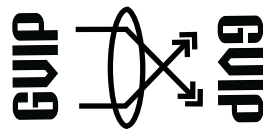


The International Congress for global Science and Technology



**ICGST International Journal on Graphics, Vision
and Image Processing (GVIP)**

**Volume (15), Issue (I)
June 2015**

**www.icgst.com
www.icgst-amc.com
www.icgst-ees.com**

**© ICGST, 2015
Delaware, USA**

GVIP Journal
ISSN Print 1687-398X
ISSN Online 1687-3998
ISSN CD-ROM 1687-4005
© ICGST 2015

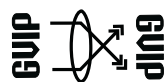
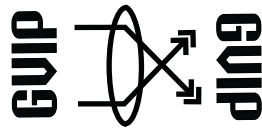


Table of Contents

Papers	Pages
P1151502359, Eman K. Elsayed and Asmaa M. Eissa, Ontology base to detect Tampering in Online Digital Images,	1--12
P1151506360, K. Srinivasa Reddy and V. Venkata Krishna and B. Eswara Reddy, Facial Recognition using the Combination of Significant non uniform Local Binary Pattern,	13--20
P1151506359, Sangita Zope-Chaudhari and Parvatham Venkatachalam and Krishna Mohan Buddhiraju, WATERMARKING SCHEMES FOR COPYRIGHT PROTECTION OF GEOSPATIAL DATA, A REVIEW ,	21--35
P1151512364, V. R. Khapli and A. S. Bhalchandra, Multistage VQ based Feature Vector for Effective CBIR,	37--43
P1151514367, Ashraf K. Helmy and Ayman H. Nasr and Safaa M. Hassan, Mutual Use of Derived Information from Radarsat-2 and ETM-8 Data for the Improvements of Feature Extraction,	45--53
P1151514368, Hind H. Zeyada and Mohammed M. Ezz and Ayman H.Nasr and Hany M. Harb, Classification of Crops using Polarimetric SAR Parameters based on Scattering Mechanism,	55-62
P1151517373, Levstein F. and Lezama J. and Maldonado C. and Penazzi, D. and Schilman M., HAMMING GRAPH BASED IMAGE COMPRESSION WITH VARIABLE THRESHOLD,	63--75



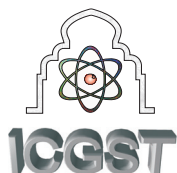
***ICGST International Journal on Graphics, Vision and Image Processing
(GVIP)***

***A publication of the International Congress for global Science and Technology -
(ICGST)***

ICGST Editor in Chief: Dr. rer. nat. Ashraf Aboshosha

www.icgst.com, www.icgst-amc.com, www.icgst-ees.com

editor@icgst.com



Ontology base to detect tampering in online digital images

Eman K. Elsayed*, Asmaa M. Eissa**

Faculty of Science (girls), Al-Azhar University, Egypt

emankaram10@hotmail.com *, asmaaeissa28@yahoo.com **

Abstract

The power of the image editor and the watermark cracking make a big challenge in determining whether the digital image is an original or doctored, determining the forgery's parts and returning to the original image.

We need to be aware that seeing does not always enough. So recently, there are many attempts to avoid and detect image forgery. A new proposed method to detect online digital images' montage was presented in this paper. The proposed method is based on semantic analysis of digital images. This proposed method starts by manipulating the original image for forgery detection process so we classify our proposed method as an active forgery detection method. Although, the proposed method has some advantages of passive forgery detection techniques as disappearing of the original image. Our methodology could be called *Semantic blind Image Forgery detection technique*. That is because the method converts the image to ontology- based. The ontology is widely used in different disciplines as a technique for representing and reasoning about domain knowledge. Using ontology comparison Algorithm for detecting image forgery guarantees the efficiency and accuracy because ontology comparison operations based on set theory and fixed point theory. Also, the proposed method is flexible to cover colored, any format and enlarge digital image. Using OWL (ontology web language) ontology engineering to detect the montage in online digital image is not only accurate for comparing based on set theory, but also important for splicing different image online databases. Also, the steganography technique is suitable to hide the universal ontology-based link inside the image.

Keywords: Ontology-based, Semantic Web, Qualitative image description QID, Steganography, active forgery detection.

Nomenclature

Symbol	Meaning
OWL	Ontology Web Language
QID	Qualitative image description
BWT	Biorthogonal Wavelet Transform
UIO	Universal Image Ontology
PSNR	Peak Signal to Noise Ratio
NC	Normalized Correlation
PMM	Pixel Mapping Method

1. Introduction

Until now completely protection the digital image from hacking, forgery or Montage is an open issue. Also, all passive detection techniques need human interpretation and not have a single straight path in terms of a detection scheme [1]. Almost all techniques in image forensics based on visual features. However, there are few works using semantic image analysis in image forensics as in reference [2]. But Yongzhen Ke et al. in [2] concentrate on the detection of abnormal semantic changes. In this paper, we propose a new framework to detect forgery image by saving the low level semantic analysis ontology link inside the image. That is to take advantage of accuracy of direct comparison between ontologies and to retrieve exactly the forgery part.

Our proposed method depends on automatic creating of Qualitative Image Description "QID" ontology and using ontology re-engineering methods as merging and comparing. That is to detect any normal or abnormal changes in the image. And also select accurately the changes. That is important to differentiate between the forgery and enhancement image. That is without any problem with adages numbers.

Also, we suggest using steganography technique. That is to guarantee detection the forgery image, without need the source image as blind or passive detection techniques.

This paper is organized as follows. Section two presents the main concepts as Quantitative Image Description (QID), Steganography, ontology base, image low level Feature Extraction Algorithms and image forgery detection techniques. The related work was



presented in section three, and then section four presents the proposed framework. The implementation found in section five. Finally, conclusion and further work are drawn in section six.

2. Main concepts

2.1 Qualitative image description (QID):

Any image describes qualitative by description the visual and spatial description of region or object, The visual feature describes the shape and color. In a shape description extracts the boundary of the object. All points of boundary calculate by the slop between a set of consecutive points. If the slope between two consecutive points and another two consecutive points are equal, this mean these points belong to the same straight segment. If the slop not equal, this mean this point belongs to curve segmentation. In the color description translate more nature color represented by the RGB (red, green, blue) into HSL (hue, saturation, lightness) to divide the interval of value according to color names.

The spatial features of any region or object represented in the topological relations and its fixed and relative orientation in two dimension space in addition to the topology relations describe the relative distance between objects. The topology situation in space description is two objects according some relations. Topological relation = (disjoint touching, completed-inside, container)

The fixed orientation description divides the space to eight orientations to obtain the oriented the object container in another object or oriented object neighbor of another object. The relative of orientation description divide the space into fifteen regions to obtain the disjoint neighbor of that object. [3]

2.2 Ontology -based:

The main aims of using an ontology is sharing, analyzing, and reusing knowledge. The artificial intelligence field contains many definitions of an ontology. One of these definitions is an explicit specification of a conceptualization. The ontology explicit description of individual, classes, attributes, relations, restrictions, rules, axioms and events. Individuals represent objects or instance in the domain. The relation is binary on individuals; i.e. the relation links two individuals together or

two classes together. Class is the concept or type of objects. Attribute is property, feature, and characteristic. The function is complex structure formed from relations between two individuals can be used them in a statement. Restrictions represented formally stated descriptions use in some assertion. The rule is statement used in logical inference to draw from assertion. An axiom is assertion in logical form. The rules and axioms are used together for comparison the all theory that the ontology describes in the domain of application. The event is changing of relations. The previous ontology definition and structure summarized from reference [4,5].

2.3 Steganography:

The Greek word steganography divided into “*Steganos*”, which mean covered or secret and “*graphy*” mean writing or drawing. Therefore, steganography means, literally, covered writing. Steganography is the art and science of hiding information such that its presence cannot be detected. [6]

Basically, the purpose of cryptography, steganography and watermark approaches are to provide secret communication. However, the steganography differentiate between cryptography and watermark. Where cryptography approach hides the contents of a secret message from attacking people, but the system is broken when the attacker can read the secret message. In steganography approach, if the system is broken the attacker need to the steganography has been used. Steganography techniques can be split into two main types, Fragile and Robust.

The watermarking approach is used to hide the small message “s” in data “m” to form a new data “M”. That is to protect the image from distortion by others, authentication and copyright protection.

Conversely, the steganography approach is used in case of small or long message “s” in data “m” to form a new data “M” to protect the image from detection and is used for the purpose of secret communication. There are a lot of algorithms for image steganography as Biorthogonal Wavelet Transform BWT, Discrete Wavelet Transform DWT, Integer Wavelet Transform IWT, AutoGuess, Discrete Cosine Transform DCT, Least Significant Bit LSB and Camouflage. There are many criteria to



compare among them as embedding efficiency, Image format,

2.4 Image low level Feature Extraction Algorithms:

A feature is defined as an interesting part of an image. The low level of features of the image is classified in overlapping groups as edges, corners, and blobs.

The feature extraction is the base of the main computer vision problems. The image low level Feature extraction Algorithms divided into two groups feature based and texture based. From our perspective in this paper, we will concentrate on only algorithms suitable for color histogram as FAST, SIFT, PCA-SEFT, F-SIFT and SURF. Reference [7] presents a comparative study among those algorithms and the authors concluded that the best algorithm generally is F-SIFT.

2.5 Image forgery detection techniques:

There are many survey articles classified image forgery detection techniques in different ways. Reference [1] classified them into the copy-move forgery, resampling, splicing geometrical inconsistency and format based. Also the authors of [1] summarized comparison among the previous techniques in the following table 1.

Technique	Advantage	Disadvantage
Robust Copy-Move Forgery	Automation possible.	Computationally expensive
Resampling Detection	Able to detect large classes of forgery.	May not in work in compressed images
Splicing Detection	Works for both Synthetic and real images	Won't Work for large size forgery as a fingerprint.
Geometrical inconsistency	Works for JPEG compressed image.	Dependent on existence of objects, enabling estimation of light.
Format Based	Frequently used format.	Results deteriorate for re-compressed forged images.

Table 1: Comparison among the forgery detection techniques

Also reference [8] presented digital image forgery detection techniques as in figure 1.

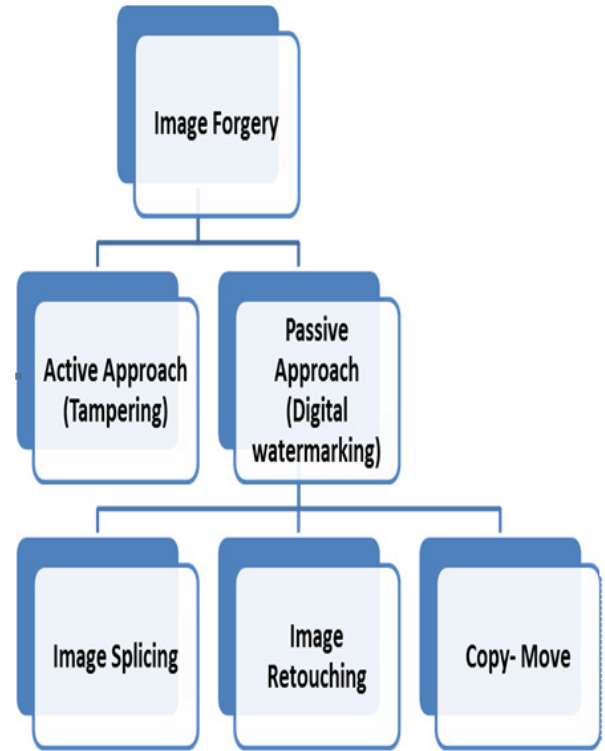


Figure 1: Classification of forgery detection techniques.

In reference [8], the image classes to set of forensic tool can be divided into five categories : 1-Pixel-based techniques detect statically anomalies of pixel level.

2-Format-based techniques measure the correlation specially compression scheme.

3-physics-based techniques to detect the anomalies in 3D between physical objects, light and the camera.

4-Camera-based techniques to detect the important factor of camera lens, sensor, or on-chip post-processing.

5- Geometric-based techniques measure the objects in the world and put them in relative position on camera

And also in reference [9] the passive image in foreign detection classes to other techniques in such that Cloning, Pixel-base, resampling, Splicing and statistical technique.

Reference [10] presented the digital image forgery detection techniques as the following classification in figure 2.



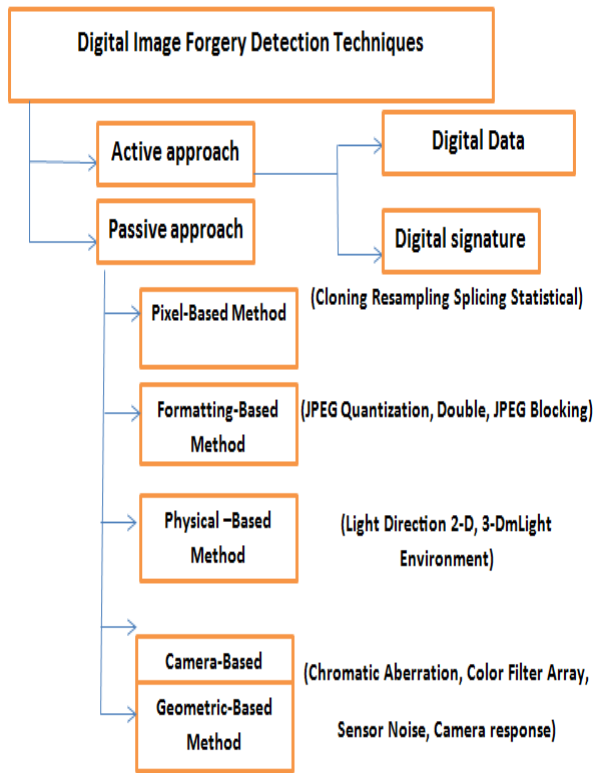


Figure2. Reference [10] digital image forgery detection techniques

3 Literature review

Reference [1] which presented passive image forgery techniques in the previous section say that Passive techniques for image forensics operate in the absence of any watermark or signature. These techniques work on the assumption that although digital forgeries may leave no visual clues that indicate tampering, they may alter the underlying statistics of an image.

Reference [2] proposed semantic framework to detect forgery image. They used high level analysis and reasoning method. But that proposal can't detect the forgery image which changes persons in a normal place.

Reference [10] proposed algorithm creates a robust secret key and embed it in the LSB of a layer of the original image to protect it against forgery. But this proposal said only if the image forgery or not. There is another proposal for detecting forgery image based generally on low level visual features as in [9],[11].

Also recently there is another branch to detect duplicated region in images as in reference [12]. Also reference [13] present novel technique for tampering detection by decomposing the image into overlapping blocks and then calculates the number of connected components in each block. Although, the proposed method in reference [13]

could detect small changes in the digital image, but it works only with the gray square image. The method proposed in reference [14] presented the general active method to detect forgery image by using code book and attached bag of geometric features as a signature to the image. The authors of this reference said that they could detect three types of image tampering as enhancing, composting and copy the movie.

There are few works using semantic image analysis in image forensics. As in reference [15] where used computer graphics and artificial intelligence commonsense reasoning techniques to detect anomalies within images.

There is another direction to secure digital image called digital watermarking as in references [16], [17]. They presented the watermark as an invisible signature inside image to protect the authentication and copyright. But it creates negative impression to viewers since the companies and photographers place the watermark in the middle of the image and sometimes it's too dark or large.

4 Proposed methodology

In this paper, we propose a new framework to detect image forgery. Our proposed mainly depends on converting image to ontology-based. Convert image to ontology-based opens the door for reasoning and expanding the knowledge about the digital image as profile for the image's owner. Our proposed method also can determine the original image parts. Using ontology comparison for detecting image forgery guarantees the efficiency and accuracy because ontology comparison operations based on set theory and fixed point theory. Our proposed method classifies into three phases.

Figure 3 presents the framework of proposed image detection based on universal image ontology based which we called *Semantic blind Image Forgery detection technique*.

The first phase starts when the owner of the digital image uploads his or her image on a certain social network, which we suggest it will be Semantic website; based on ontology based. Then the system automatically starts create Universal Image Ontology (UIO). This ontology integrates the qualitative ontology of the image; the qualitative ontology integrates all knowledge about visual and spatial features



which described in Qualitative image description QID analysis. Also UIO have owner profile. Note, this is the first version of the UIO, where we can connect it to any useful ontology based.

In the second phase the admin hides the ontology link inside the image. That is by any steganography technique.

In our proposed method, we hide only ontology link. That is for more protection. Because in our semantic web site the links are more secure and we always assume that the attacker knows that there is hidden information inside stego-image.

In this phase, we use any steganography Algorithm as Biorthogonal Wavelet Transform BWT or AutGuess Algorithm. The choice depends on the image types and the degree of website security.

The third phase starts when someone upload the image to the Semantic social network. Then admin checks if there is an ontology link inside the new image. That is by un-hide method. This phase is classified into two classes:

If the admin doesn't find UIO link, then he would run procedure in phase one and gives a permission to upload the image.

But if he found UIO link, then he would run the ontology comparison algorithm. This algorithm compares between the founded UIO and the UIO of the uploaded image. May be the visual similarity algorithms are accurate and easy. But they need a huge memory to save images and if the original image deleted, then the direct image comparison is impossible. So we use the ontology comparison algorithm.

The ontology comparison Algorithm compares the structure of the ontologies not the textual ontologies. First, The algorithm runs the extensible set of heuristic matchers and the results are combined by fixed point algorithm.

The ontology comparison Algorithm runs in two steps: First, the alignment step to detect the difference between frames of two ontologies and find any refactor is changed. Where the *structure* of the comparison between two ontologies is defined as the following

$$\text{Diff}(\text{UIO1}, \text{UIO2}) = \{(f_i, f_j): (f_i \in \text{UIO1} \vee f_i = \text{null}) \wedge (f_j \in \text{UIO2} \vee f_j = \text{null})\}$$

Where, $f_i == f_j$ i.e. f_i matches f_j

$f_i \neq f_j$ i.e. f_i does not match f_j

$f_i \neq f_j$ when $f_i = \text{null} \vee f_j = \text{null}$

Second, the explanation step which organizes the axioms changed which result in the alignment phase, then displays the data by understanding manner that would be more readable. The output of this phase can determine exactly if the two UIO identical, differing in main objects or different in secondary features.

5 Case study

In this section, we present, how can we experiment the method in section four. In our proposed method OWL ontology language is more suitable for online application as semantic web. So we used Protégé 4.3 Platform and its library plugin as comparing and merging tool. Where it is free, open source OWL ontology platform and a framework for building intelligent system [18]. Also, steganography algorithms are found in [19] where Steganography Studio software is a tool to learn, use and analyze key steganographic algorithms. The proposed method has been evaluated using different data sets which contain different types of images as in reference [20] database. The Shampoo 12 editor [21] used for tampering the training set of 60 images.

The first step is image feature extraction. Where we extract color image qualitative description as in figure 4. After discussing the image feature extraction in references [22, 7] we select image features extraction algorithm F-SIFT.

The flexibility of our method is an inflexible change in algorithm in each phase separately.

Then the method converts automatically the Qualitative Image Description file to Qualitative ontology, as in figure 5.

Also, the method converts the owner profile to ontology and merge the two ontologies to generate the universal image ontology UIO.

That is by Prompt tool from Protégé platform. The profile of the image is shown in figure 6. This ontology has information about an owner who created the image, modified date and time, Image format, image resolution and its size.

Then the second step is hiding the UIO link inside the image. Where, the method starts a steganography model to hide the UIO link as a message inside the image as a carrier or cover an object. The general model is as in figure 7. We can hide the ontology inside the image, but



hiding ontology's link is more secure. Also the ontology file's size is heavier than its link to the image. The standard range of link's size is nearly 8 KB.

In case of using a camouflage algorithm, it appends the link into the end of the image. It is quickly cracked, but not make any noise on the image.

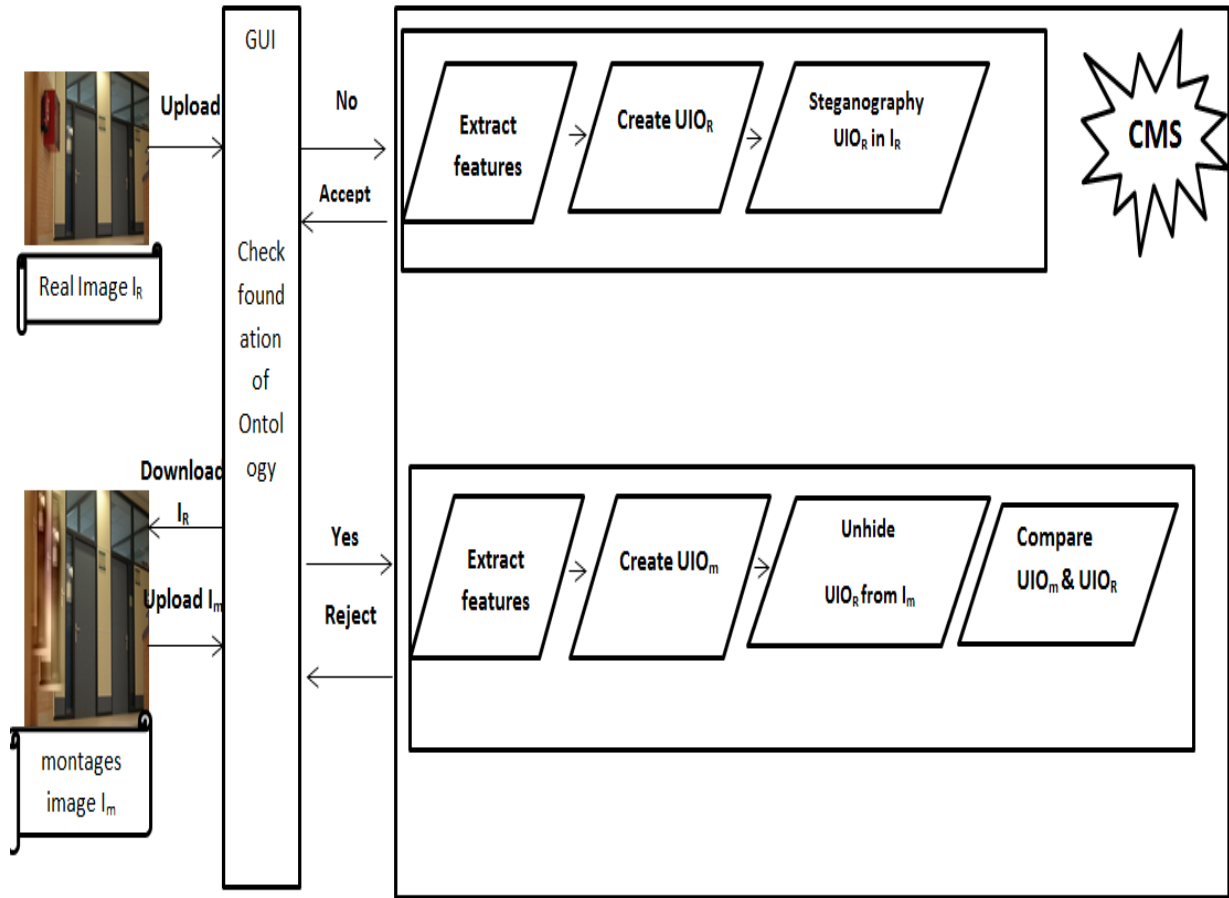


Figure 3. Semantic blind Image Forgery detection technique.

In case of using an OutGuess algorithm, the embedding efficiency average - the average number of bits embedded per embedding change - is 0.96 when a maximum message capacity is less than 6.5% that is satisfied in our case study. In case of using Biorthogonal Wavelet Transform BWT Algorithm the Peak Signal to Noise Ratio PSNR is nearly 35.95 and the Normalized Correlation NC is 97.19%. But if our achievement is the indetection against some well known steganalysis attacks, then we can use PPM Pixel Mapping Method [23].

Now for testing the framework, we download the image from the semantic website and montage the image (I_m), when we try to re-upload the image after montage. Then the following steps run as in figures 8a,8b,8c and

the system sends the reject message. That is after having detailed report about the image.

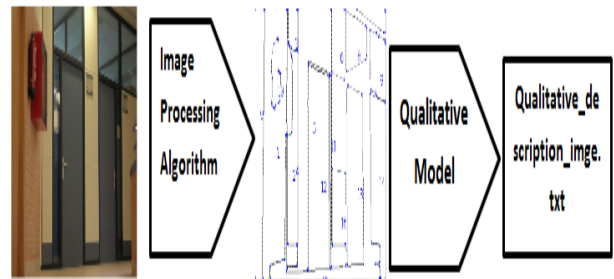


Figure4. Feature extraction process

After applying several types of comparison algorithms, we found that using ontology's comparison explain the difference in more details as COMP algorithm [24] and PromptDiff algorithm [25]. So we compare



between two OWL ontologies automatically without any intervention. We use the version comparison algorithm for analysis the difference between the two versions of ontology in the same domain. The COMP algorithm loosely depends on PROMPTDiff Algorithm which used for comparing ontologies. Then we use the “Diff” tool based on using semantic of the definitions and axioms of OWL in order to give the admin detailed report about the changes. That is to differ between forgery and only enhancement images. That is shown in figure 9.

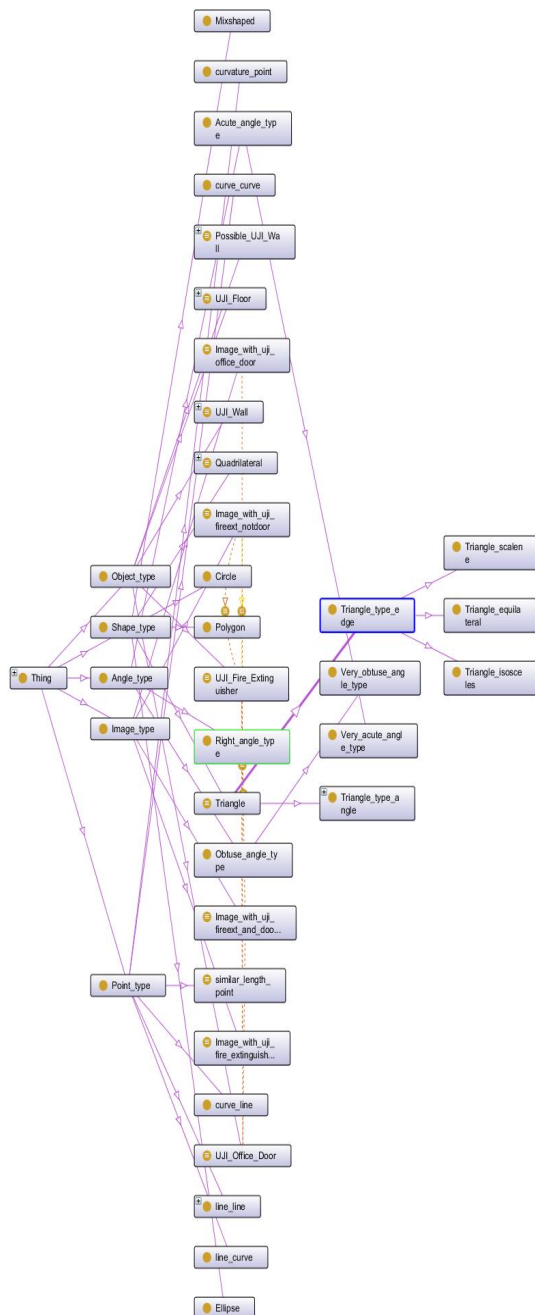


Figure5. IDO Image description ontology

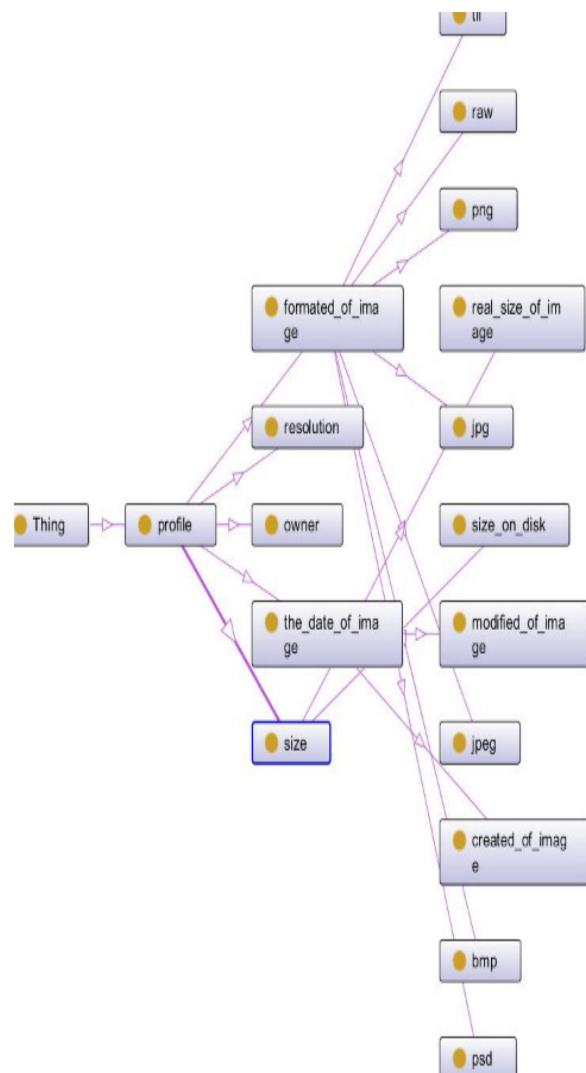


Figure 6. Image profile ontology

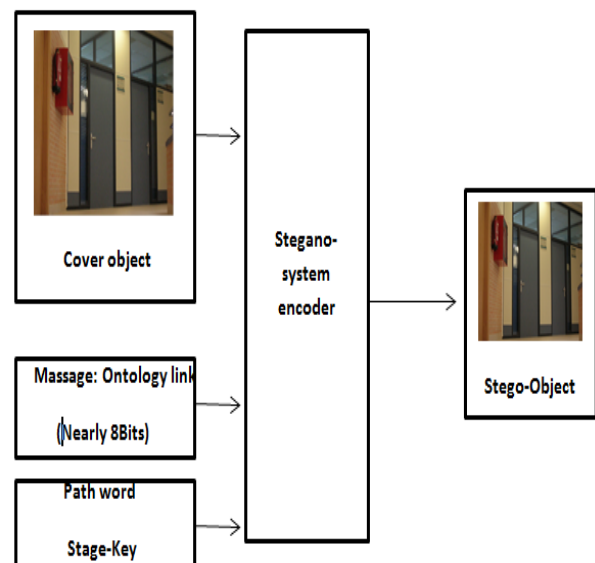
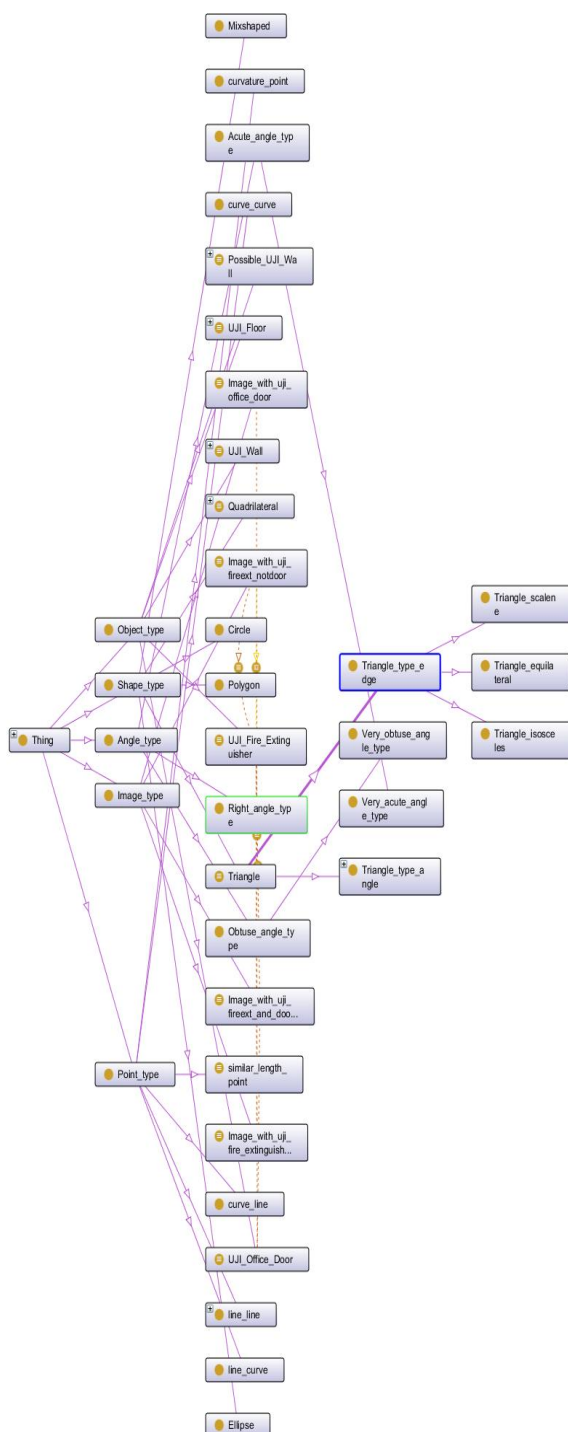


Figure 7. Steganography model





Figure8a. Montages image Im



Figur8b. Unhide the ontology link from I_m

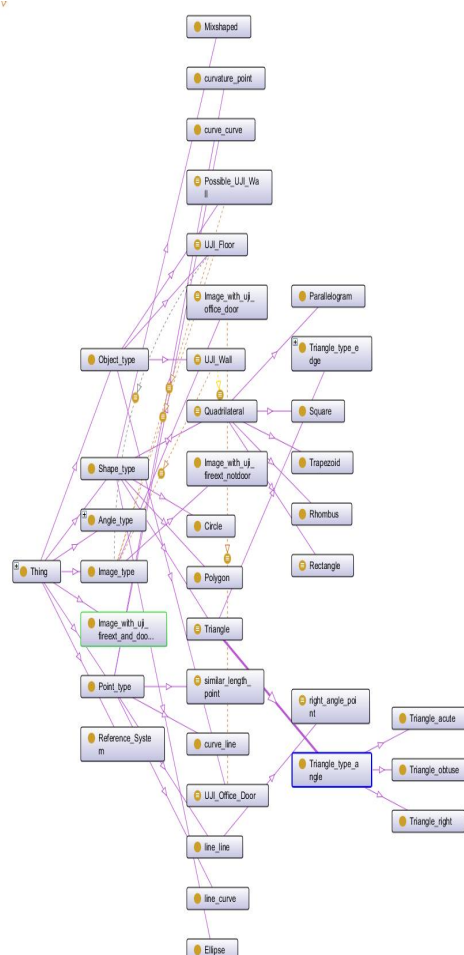


Figure8c. Automatically Create UIO for I_m

“PromptDiff “uses a Frame-like API where an ontology is viewed as a container of property values for names and anonymous entities. To compare between two ontologies “PromptDiff” Algorithm has two phases. **First**, the alignment phase where it detects the difference between signatures of two ontologies and finds any refactor is changed. Prompt Diff Algorithm in the alignment phase classifies the ontology’s changes into five different operations:

1. Deleting the existing object (frame) from UIO of the original image.
2. Adding to the existing object (frame) to UIO of the Forgery image.
3. Merging two objects (frames) from the original image have been combined in the UIO of the Forgery image which means change in objects.
4. Splitting an object from UIO of the original image into two objects in the forgery image.
5. Mapping the objects exist in both ontologies and the previous four operations don't apply. But Also, there are three levels of



Map Operation its (unchanged, changed, isomorphic)

5.1 Unchanged operation means identical objects (frames).

5.2 Changed operation means the frames have slots and facet values are not identical.

5.3 Isomorphic operation means the frame slots and facet values are match of each other, but not identical matches.

Ex. The frame referenced by one of the slots may have changed between ontologies.

The result of the alignment phase is the tuple

$\langle (f_i, f_j), \text{renamevalue}, \text{operation value}, \text{maplevel} \rangle$

Where, (f_i, f_j) is the result of structure Diff (UIO1, UIO2)

Rename value is True if there is the same frame names for f_i, f_j & False otherwise.

Operationvalue $\in \{add, delete, split, merge, map\}$


map level $\in \{change, unchange, isomorphic\}$

Second, the explanation phase where it organizes the axioms changed which result in the alignment phase, then display the data by understanding manner that would be more readable.

The promptdiff Algorithm running time is $O(M)\log(M)$ time, where $M = \max(n, m)$; n, m are the frames number of UIO1 and UIO2 respectively.

Also in the small modification as the change in figures (10a, 10b) which we can't see but ontology can do.

Using the ontology's comparison tool gives admin details report about the effect of changing color on certain objects. The comparison report depends on the ontology based, where the figure (11) presents a semantic comparison report for figures (10a, 10b).



Ontology Differences

X

Find

Modified: UJI_Office_Door	Source and target entities aligned because they have the same IRI.		
	Description	Baseline Axiom	New Axiom
	Definition changed	UJI_Office_Door EquivalentTo Object_type and ((has_colour value blue) or (has_colour value grey)) and (has_shape some Quadrilateral) and (is completely inside some Image type)	UJI_Office_Door EquivalentTo Object_type and ((has_colour value blue) or (has_colour value dark_grey) or (has_colour value grey)) and (has_shape some Quadrilateral) and

Figure9. Semantic Comparison Report



Figure 10a: the image without change



Figure 10b: the image with normal change



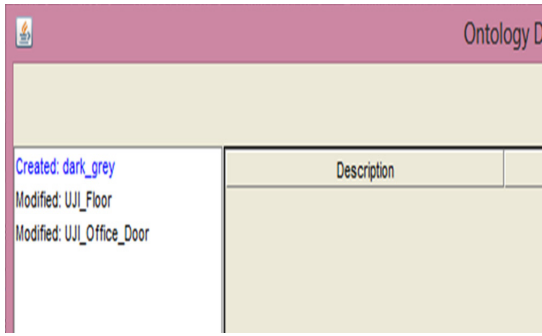


Figure 11: Semantic Comparison Report

There are image comparison algorithms as “Resemble.js”[26] and “comp”[27] (short for complimentary). Where, “Resemble.js” can be used for any image analysis and comparison requirement you might have in the browser. However, it has been designed and built for use by the Phantom JS. Figure 12 displays The image’s comparison result between I_m and I_R .

In this part of our proposed method, The robustness of using ontology comparison instead of image comparison appear in the following case. If we change the image’s enlarge, then the Image comparison tool as “resemble tool” can’t determine the change. That is as in figure 13a. But in case of using the ontology comparison tool we can catch small change in size. Also in our method we can catch a change in owner profile. Our proposed method compares without knowing where the real image uploaded. Also reasoning as a high level image analysis could be applied in ontology analysis case.

Now, to measure the efficiency and robustness of the method we use the precision rate and the recall rate.

Where the precision rate is defined as the ratio of correctly detected parts to correctly detect parts plus false positive. False positive are these parts which are original, but have been detected as tampered parts.

$$\text{Precision} = \frac{\text{Correctly detected parts}}{\text{Correctly detected parts} + \text{false positives}} \times 100$$

The Recall rate is defined as the ratio of correctly detected parts to correctly detect parts plus false negatives. False negatives are these parts, which are actually tampered parts, but have not been detected.

$$\text{Recall} = \frac{\text{Correctly detected parts}}{\text{Correctly detected parts} + \text{False Negatives}} \times 100$$

The data set consists of 60 different image types. Table 2 displays the precision and recall rates for different image types.

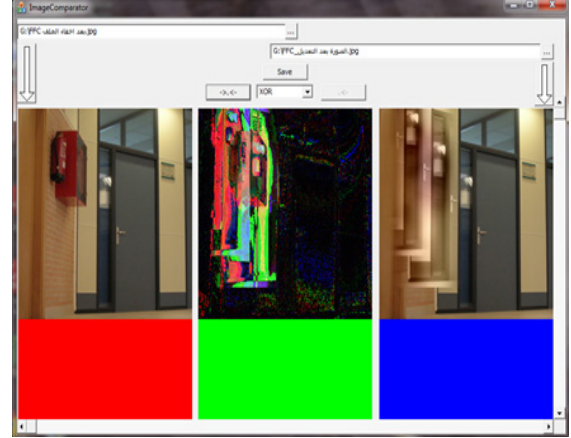


Figure 12 image's comparison by image comparator

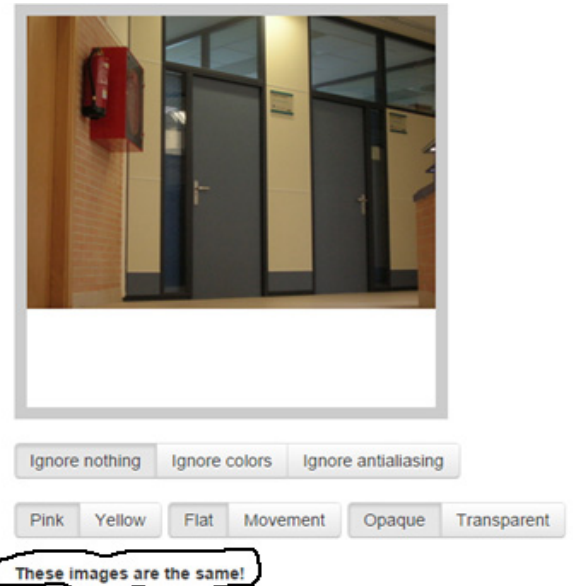


Figure 13a: compare by “resemble” image comparison tool

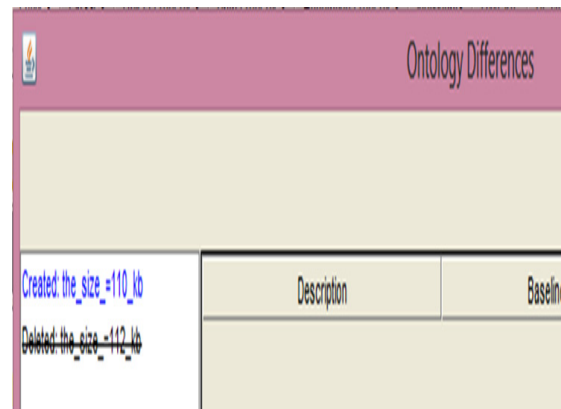


Figure 13b: compare by “Ontology Comparison tool”





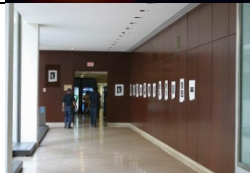

Forgery image example	Type of modification	Precision	Recall
	Enhance color	98.7	98.2
	Add an object	97.5	97.4
	Copy-move	30.2	30.1
	Delete object	97.8	97.2

Table 2: display the accuracy by measuring precision and recall of the data set

6 Conclusion and future work

Until now, no methods achieved 100% accuracy or robustness against forgery. The proposed method “*Semantic blind Image Forgery detection*” was presented in this paper detecting the tampering in online digital images. It is flexible to cover any image format, colored and enlarge the image. The relation between the run time of comparing step and the size of the universal image ontologies is linear and equal $O(M)\log(M)$.

The advantages of our proposed framework are: online storage place is miniature; each phase is flexible updating separately, and also change QID ontology to the other ontology type of image support the reusability. The proposed method bases on image semantic analysis. That was by automatic creating for image universal ontology based. The universal ontology must have all knowledge about image as its description Qualitative image description QID and owner profile.

Also, we use steganography technique to hide the universal ontology link inside the image. This hiding link increases the image weight just eight KB only.

The degree of security depends on the rules of the website. Also table 2 said that our proposed method is accurate in high percentage in different cases except copy-move case. So the new version of the ontology must be has a solution for this type of forgery and also able to change the steganography algorithm depended on the type of image and degree of security.

Finally, we can consider that *Semantic blind Image Forgery detection* is the first version to enrich and ameliorate image comparison methods to detect forgery. That is by integrating proposed ontology for extract low level features with ontology for high level features.

From the fact that ontology base is content driven method so our proposal are used for certain type of place analysis. In the future, we can create many general ontology bases for different place types as indoor place, animals, or nature. Also in the future we will apply ontology comparison algorithms to be suitable for OWL2.

7 References

- [1] V. Archana, S. Mirei, N. Dhaka, and P. Prey, Catalogue of Digital Image Forgery Detection Techniques, an Overview, Proc. Of Int. Conf. On Advances in Information Technology and Mobile Communication, Elsevier, 2013.
- [2] Y. Weidong, M. Fan Qin and J. Shang, Image forgery detecting based on semantics, International Journal of hybrid information technology, Vol. 7, pp. 109-124,2014.
- [3] Z. Falomir, E. Jimenez-Ruizo, M. Teresa Esrig and L. Museros, Description images using qualitative models and description logics, Spatial cognition and computation Taylor, Francis, 2010.
- [4] X. Zhang., B.Hu, Xu Ma, F. Moore., J. Chen, Ontology driven decision support for the diagnosis of mild cognitive impairment, Computer Methods and Programs in Biomedicine, Volume 113, Issue 3, pp. 781-791, Science Direct. March 2014.
- [5] P. Ordoñez et al, Advancing Information Management through Semantic Web Concepts and Ontologies, book, IGI Global Idea Group Inc. 2012.
- [6] T. Morkel, J. Eloff, M. Oliviers,”An overview of image steganography”, Information and computer security architecture research group, 2011.
- [7] M. El-gayar, H.Soliman, N. Meky, A comparative study of image low level feature extraction algorithms, Egyptian Informatics Journal Vol. 14, PP. 175–181,2013, www.elsevier.com/locate/eij.
- [8] P. Gomase1, N. Wankhade, Advanced digital image forgery detection: a review, IOSR journal of computer science p. 80-83 (IOSR-JCE) 2014.



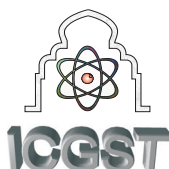
- H. Farid, "Image forgery detecting survey", IEEE signal processing magazine March 2009.
- [9] A. Tafti, M. Malakooti, RGB Digital Image Forgery Detection Using Singular Value Decomposition and One Dimensional Cellular Automata, NCM2012 8th IEEE international conference on computing and advanced information management Soul, Korea 2012.
- [10] N. Kadam, and D. Balk, Forgery detection in digital image, IJAET International Journal of Advances in Engineering & Technology, May 2013.
- [11] P. Mishra, N. Mishra, S. Sharma, and R. Patel, Region Duplication Forgery Detection Technique Based on SURF and HAC, The Scientific World Journal Vol. 2013, <http://dx.doi.org/10.1155/2013/267691>
- [12] N. Tiwari, N. Hemrajani & M. Kumar, A novel technique for tampering detection, International Journal of Computer Science, Engineering and Information Technology Research (IJCEITR) ISSN 2249-6831 Vol. 3, Issue 3, 341-346, 2013
- [13] E. Agnes, S. Devi Mahalakshmi, K. Vijayalakshmi, A Forensic Method for Detecting Image Forgery Using Codebook, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 3, pp. 1-5, March 2013.
- [14] S. Lee, D. A. Shamma and B. Gooch, "Detecting False Captioning using Commonsense Reasoning", Digital Investigation, vol. 3, no. 1, pp. 65-70, 2006.
- [15] J. Lichtenauer, I. Setyawan, T. Kalker and R. Lagendijk, Exhaustive geometrical search and the false positive watermark, Information and communication theory group, 2003.
- [16] M. Thapa and S. Sood, On Secure Digital Image Watermarking Techniques, Journal of Information Security, Vol. 2 No. 4, pp. 169-184, 2011.
- [17] Protégé platform <http://protege.stanford.edu/> .
- [18] Steganography Studio software http://steganography_studio.en.softonic.com/ .
- [19] D. Tralic, I. Zupancic, S. Grgic, "CoMoFoD - New Database. 55th International Symposium ELMAR-2013, pp. 49-54, September 2013 <http://www.vcl.fer.hr/comofod>
- [20] Ashampoo photo commander 12 Software <https://www.ashampoo.com/en/usd/fdl>
- [21] S. Kumar, E. Chauhan, A survey of feature extraction techniques for color images, International Journal of Scientific & Engineering Research IJSER, Volume 5, Issue 9, 2014 <http://www.ijser.org>.
- [22] S. Bhattacharyya, B. hattacharyya, A. Khan, I. Banerjee, A Robust Image Steganography method Using PMM in Bit Plane Domain, International Journal of Computer, Information, Systems and Control Engineering Vol:8 No:9, 2014
- [23] N. Taleb, An Algorithm and a Tool for Comparing Ontology Versions, International Journal of Software Engineering and Its Applications Vol. 7, No. 3, May, 2013
- N. Noy & M. Musen, PromptDiff: A Fixed-Point Algorithm for Comparing Ontology Versions. In the Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-2002), Available as SMI technical report SMI-2002-0927, 2002.
- [24] <http://webscripts.softpedia.com/script/Image-Galleries/Image-Tools/Resemble-js-78435.htm>
- [25] <http://www.softpedia.com/get/Multimedia/Graphics/Graphic-Others/Image-Compare-Timothy.shtml>



Eman Karam Elsayed, assist. Prof. Computer science, Alazhar university, Master of computer science, Cairo University 1999, Bachelor of Science, mathematics and computer science Department, Cairo University 1994. I Published thirty one papers until 2014 in data mining, Ontology engineering, e-learning and software engineering. I also published two books in Formal methods and event B on Amazon database. I am a member of Egyptian mathematical society and Intelligent computer and information systems society.

Asmaa Mahmood Eissa postgraduate Computer science student in Faculty of Science, Alazhar University since 2008.





Facial Recognition using the combination of significant non uniform local binary patterns

K. Srinivasa Reddy¹ V. Venkata Krishna² B. Eswara Reddy³

¹ Associate Professor, Dept. of CSE, BVRIT Hyderabad College of Engineering for Women, Hyderabad, T.S., India.

Email: kondasreenu@gmail.com

² Principal, G.I.E.T, JNTUK, Kakinada, Rajahmundry, A.P., India.

Email: vakula_krishna@yahoo.co.in

³ Professor, Dept. of CSE, JNTUA, Ananthapuram, A.P., India.

Email: eswarcsejntu@gmail.com

Abstract

Efficient face recognition is one of the long standing problems of computer vision. Recently, Local Binary Pattern (LBP) has proven to be an effective descriptor for object recognition in general and face recognition. This paper presents an efficient algorithm for face recognition by deriving a new set of stable transitions on LBP for selecting Significant Non Uniform LBP's (SNULBP). The proposed SNULBP's are stable, because it considered the transitions from two or more consecutive zeros to two or more consecutive ones. The proposed Significant NULBP (SNULBP) along with Uniform LBP (ULBP) features improved facial texture recognition rate. The performance of the proposed scheme is validated using complex facial datasets, namely Yale, Indian and American Telephone and Telegraph Company (AT&T) Olivetti Research Laboratory (ORL) with various facial expressions.

Keywords: Significant NULBP (SNULBP), Uniform LBP, Transitions on local binary pattern, Face recognition.

1. Introduction

Face recognition has attracted much research effort for many years due to ever growing applications. Further a lot of attention has been received in face recognition research due to its theoretical challenges as well as its many applications. The researchers in face recognition, age classification based on human faces and facial expression recognition designed methods to work best with, well illuminated, well aligned and frontal pose face images. The researchers developed many successful face recognition systems [1,4,16,17, 21,30,36], and Zhao et al. [45] provided a good survey. The early work on face recognition extracted global features based on sub space methods.

The examples for this are Eigen and Fisher faces methods [4, 36]. These methods projected the whole face into a linear subspace to acquire or identify face variations. These methods perform well under ideal circumstances. This trend is changed completely and most of the face recognition algorithms of today are based on local features. This is because they are simple and robust in encoding the relevant local traits: to harness the local information the method [6] used vector quantized local pixels, the other methods uses a battery of spatially localized gabor filters in multi layer frame work for face recognition [7,20], the most popular methods that are based on Local Binary Patterns [LBP] extracted from intensity images uses a histogram of local pattern features [1,4,40], in other methods [25,38,43] local features are extracted from image orientation. There is a need to develop robust face recognition methods that works well under a variety of situation such as illumination and pose variations. In some other applications especially when working with surveillance cameras, automatic tagging, and human robot interaction, however, it is not possible to meet these conditions. To address the above, some researchers derived methods on unconstrained face images [8,32,39,41]. Researchers working with unconstrained facial images used SIFT models [5,15], local appearance descriptors such as Gabor jets [35,47], wavelet transforms [18], histograms of Local Binary Patterns [27], Speeded Up Robust Features (SURF) [3], Histogram of Oriented Gradients (HOG) [2]. Several alternative models are used in the literature to represent faces. Linear models [4,36], linear SVM [6,11], different similarity metrics are also used to compare and evaluate faces, the popular among them are distance based metrics such as Euclidian distance [11,40], angle based cosine similarity [24,29,35,38] and the methods based on chi-square distance methods [1]. Researchers are also



adopted methods by dividing the face into regions and evaluated local traits for effective face recognition [1,16,34]. These methods achieved good recognition rates for the small regions instead of big regions. LBP [13,14,26,27,46] are widely used in the literature for many image processing applications because of their local computationally efficient nature and robustness in representing local features and illumination variation. One of the disadvantage of LBP's is considering the huge number of Non ULBP's under one label called miscellaneous and by which some information may be lost. The present paper addresses this.

The remainder of the paper is organized as follows. The section (2) describes the LBP methodology, section (3) presents the proposed methodology, section (4) presents the results and discussion and section (5) presents the conclusions.

2. Local Binary Pattern

The Local Binary Pattern (LBP) was introduced by Ojala et al [26] in 1996. LBP is simple, computationally efficient, robust, and derives local attributes efficiently. With these features, many researchers started working with LBP in various domains and especially in face recognition [1,31,32,39]. The LBP is a powerful tool to describe the local attributes of a texture. In the LBP the grey level image is converted into binary by taking the central pixel value as a threshold and this grey level value is compared with its neighborhood values. The resulting binary valued image is treated as a local descriptor. The basic LBP was initially derived on a 3*3 neighborhood. This LBP operator can also be represented with different variation of (P,R) where P represents the number of neighborhood pixels and R is the Radius. By this the basic LBP operator is represented as (8,1). The 8-bit binary representation or 8-neighboring pixels on a 3*3 neighborhood or (8,1) derives a LBP code that ranges from 0 to 255. The LBP operator takes the following form as given in equation 1.

$$LBP(8,1) = \sum_{n=0}^7 2^n S(P_c - P_n) \quad (1)$$

Where 'n' runs over the 8 neighbors (0 to 7) of the central pixel C, P_c and P_n are the grey level intensities at c and n and $S(u)$ will be 1 if $u \geq 0$ and 0 otherwise. The LBP encoding process on a 3*3 neighborhood i.e. (8,1) is given in fig.1.

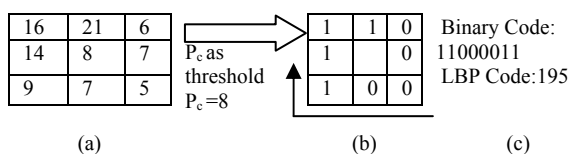


Figure.1. Encoding of basic LBP operator

3. Derivation of SNULBP

Many researchers worked on Uniform Local Binary Pattern (ULBP) and Non Uniform Local Binary Pattern (NULBP) and derived many conclusions. An LBP is uniform if it contains at most one - zero to one and one - one to zero transition in a circular manner. For example 11111111 (0 transitions), 00000001 (2 transitions) are uniform, where as 11001100 (4 transitions), and 01011111 (5 transitions), 10101100 (6 transitions), 01010101 (8 transitions) are not uniform. Some of the researchers [1,14,19,27,28,44] considered only ULBP's for classification, recognition and for solving other problems because of the following reasons. a) ULBP's are treated as the fundamental properties of texture image. b) 80 to 85% of the texture images contain only ULBP's. c) There are 192 NULBP's and treating them as miscellaneous will reduce lot of dimensionality without losing much of the texture content. The other group of researchers [9,10,13,22,23,37,46] also considered a part or few of NULBP's along with ULBP's and proved that this combination yielded a better or a little progress than by considering only ULBP's. From this one can understand that ULBP's can be treated as the fundamental properties of the texture image, however considering them only may lose some basic information. Therefore it is better to consider a sub set of NULBP's.

Different authors considered different sets of NULBP's. In our previous work we have defined Prominent LBP (PLBP) to solve the above problem [37]. The PLBP contains the combination of prominent ULBP's (not all ULBP's) and prominent NULBP's. The PLBP contains a new set of transitions that are completely different from the formation of ULBP's. The PLBP considered the transitions that occurs after two or more consecutive zeros immediately followed by two or more consecutive ones and vice versa in a circular manner. 92 different LBP forms the PLBP on a 3 x 3 neighborhood with a radius of one. Out of these 40 PLBP's belongs to ULBP's and 52 belong to NULBP's. Based on the above new transition rule the PLBP treats 18 ULBP's and 146 NULBP's on one label called "miscellaneous". The present paper considered Smallest PLBP (SPLBP) by using $PLBP \cap ULBP$. The SPLBP contains 40 ULBP's only and it treats the remaining 216 LBP's (which contains 18 ULBP's and 198 NULBP's) as miscellaneous set.

From the above discussion it is evident that the major problem is how to select a subset from NULBP's to improve the overall performance and to reduce overall dimensionality. For this the present paper derived Significant NULBP (SNULBP). The SNULBP is a subset of NULBP. They contain transitions from two or more zeros to two or more ones and vice versa is not true. The transitions are measured in a circular manner. No ULBP will have such transitions. Table 1 illustrates why the ULBP's does not fall into SNULBP. Table 2 illustrates why



certain NULBP's does not fall into the group of SNULBP. Table 3 illustrates why some of the NULBP's fall into the group of SNULBP.

Table 1. ULBP that does not fall into SNULBP

ULBP CODE	LBP	REASON
12	00001100	It is having the transitions from 00 to 11 and also 11 to 00. Therefore it is not a SNULBP.
15	00001111	It is having the transitions from 00 to 11 and also 11 to 00. Therefore it is not a SNULBP.
32	00100000	It is not having transition from 00 to 11 at all.
135	10000111	It is having the transitions from 00 to 11 and also 11 to 00. Therefore it is not a SNULBP.
227	11100011	It is having the transitions from 00 to 11 and also 11 to 00. Therefore it is not a SNULBP.

Table 2. NULBP'S that does not fall into SNULBP

NULBP CODE	LBP	REASON
17	00010001	It is not having transition from 00 to 11 at all.
35	00100011	It is having the transitions from 00 to 11 and also 11 to 00 in circular manner. Therefore it is not a SNULBP.
85	01010101	It is not having transition from 00 to 11 at all.
175	10101111	It is not having transition from 00 to 11 at all.
197	11000101	It is not having the transition from 00 to 11 at all.

Table 3. NULBP'S that fall into SNULBP

NULBP CODE	LBP	REASON
13	00001101	It is having transition from 00 to 11 and not having transition from 11 to 00.
67	01000011	It is having transition from 00 to 11 and not having transition from 11 to 00.
104	01101000	It is having transition from 00 to 11 in circular manner and not having transition from 11 to 00.
158	10011110	It is having transition from 00 to 11 and not having transition from 11 to 00.
232	11101000	It is having transition from 00 to 11 in circular manner and not having transition from 11 to 00.

The derived SNULBP's are stable because we are considering the transitions that occur from two or more consecutive zero's to two or more consecutive one's only, instead of zero to one or vice versa. For efficient face recognition the present paper combined the derived SNULBP's with ULBP, PLBP and SPLBP using union operation only. 90 different LBP's are formed out of 256 by SNULBP U ULBP, in the same way there will be 124 and 72 different

LBP's by using union operation in between SNULBP U PLBP and SNULBP U SPLBP respectively.

The SNULBP U PLBP contains 40 ULBP's and 84 NULBP's. The SNULBP U SPLBP contains 40 ULBP's and 32 NULBP's only. For efficient face recognition the present paper evaluated various features based on histograms of ULBP, PLBP, SNULBP, SNULBP U ULBP, SNULBP U PLBP and SNULBP U SPLBP with different (P, R) (where P corresponds to the number of neighboring pixels considered on a circle of radius of R) on each individual facial image and placed in training database. In the similar way the above histograms are evaluated for test facial image and the face recognition is evaluated based on Chi-square distance [1] method as given in equation 3.

$$R(d, t) = \min \left(\sum_{i=1}^n ((d_i - t_i)^2 / (d_i + t_i)) / 2 \right) - (3)$$

Where d, t are two image features (histogram vectors) and R(d,t) is the histogram distance for recognition.

Table 4: Recognition rate by considering new test images (not part of training database) for Yale database: Initial case.

(P,R)	ULBP	PLBP	SNULBP	SNULBP U ULBP	SNULBP U PLBP	SNULBP U SPLBP
(8,1)	64.44	66.67	15.56	66.67	71.11	66.67
(8,2)	66.67	68.89	26.67	68.89	72.22	68.89
(8,3)	68.89	68.89	37.78	73.33	73.33	71.11
(8,4)	66.67	71.11	48.89	71.11	75.56	68.89
(16,1)	75.56	73.33	11.11	76.67	75.56	75.56
(16,2)	75.56	75.56	22.22	77.47	76.67	75.56
(16,3)	68.89	75.77	26.67	75.56	77.47	76.11
(16,4)	71.11	76.67	28.86	78.67	78.67	71.11
Average	69.72	72.11	27.22	73.55	75.07	71.74

4. Results and discussion

The present paper considered facial images from Yale data base [42], Indian database [12] and AT&T ORL database [33] for face recognition. For training data base the present paper considered 120, 472 and 320 facial images with different facial expressions from Yale data base [42], Indian database [12] and AT&T ORL database [33] respectively for face recognition. This indicates the present paper considered only 75% to 80% of the facial images for training set. The present paper evaluated the above texture descriptors on different (P, R). For efficient face recognition the present paper evaluated Chi-square distance [1] as given in equation 3. The present approach considered the remaining leftover facial images of the above three databases, which are not considered for training set, as test images in the initial case. The recognition rate for the above three databases for initial case are shown in tables 4, 6 and 8. In the second case the present paper experimented by considering the test



images as a combination of new and training database images and the face recognition rates are given in tables 5, 7 and 9 for the above databases.

Table 5: Face recognition rate for Yale database for case 2 (The combination of new and training set as test images).

(P,R)	ULBP	PLBP	SNULBP	SNULBP U ULBP	SNULBP U PLBP	SNULBP U SPLBP
(8,1)	80.97	82.08	16.78	82.65	84.56	81.15
(8,2)	82.08	84.20	28.87	84.13	86.65	82.25
(8,3)	83.19	85.45	41.11	86.85	87.88	83.35
(8,4)	82.08	86.54	52.33	87.88	90.55	83.75
(16,1)	86.53	85.42	14.56	87.88	88.85	83.15
(16,2)	86.53	86.53	25.45	88.85	89.66	85.55
(16,3)	83.19	88.66	28.94	88.64	91.65	85.75
(16,4)	84.31	88.77	32.33	89.97	92.33	86.41
Average	83.61	85.96	30.05	87.11	89.02	83.92

Table 6: Recognition rate by considering new test images (not part of training database) for Indian database: Initial case.

(P,R)	ULBP	PLBP	SNULBP	SNULBP U ULBP	SNULBP U PLBP	SNULBP U SPLBP
(8,1)	87.57	88.14	18.64	91.15	92.55	87.57
(8,2)	87.57	88.70	30.51	91.55	92.70	88.11
(8,3)	89.27	88.70	46.33	93.15	93.55	88.35
(8,4)	90.40	91.53	62.71	94.35	94.55	89.83
(16,1)	92.66	93.44	15.65	95.45	96.55	91.53
(16,2)	91.53	94.92	27.12	95.15	97.66	94.35
(16,3)	94.92	95.65	33.90	96.00	98.15	94.92
(16,4)	94.35	96.77	34.46	96.15	98.55	94.35
Average	91.03	92.23	33.67	94.12	95.53	91.13

Table 7: Face recognition rate for Indian data base for case 2 (The combination of new and training set as test images).

(P,R)	ULBP	PLBP	SNULBP	SNULBP U ULBP	SNULBP U PLBP	SNULBP U SPLBP
(8,1)	92.54	92.82	22.45	94.33	95.15	92.54
(8,2)	92.54	93.65	31.51	94.53	95.75	92.81
(8,3)	93.38	94.12	47.56	95.33	96.45	92.93
(8,4)	93.95	94.52	64.21	95.93	97.15	93.67
(16,1)	95.08	95.47	18.65	96.55	97.25	94.51
(16,2)	94.52	96.21	29.11	97.25	97.55	95.93
(16,3)	96.21	96.58	34.25	97.55	98.45	96.21
(16,4)	95.93	97.14	35.56	98.00	98.75	95.93
Average	94.27	95.06	35.41	96.18	97.06	94.31

From the tables 4, 5, 6, 7, 8, 9 the following factors are noted.

- 1) The face recognition rate for Indian and AT&T ORL databases are high when compared to Yale database for cas

- 2) e1 and case 2. This is shown in the form of graph of fig.5. The face recognition rate for Indian and AT&T ORL databases are high when compared to Yale database for case1 and case 2. This is shown in the form of graph of fig.5.
- 3) The face recognition rate is high for (P_2, R) when compared to (P_1, R) where $P_2 > P_1$ i.e. the considered neighborhood points are more, for the same radius. This is evident for all databases on all proposed methods. The reason for this is the number of NULBP's will increase as we increase the number of neighboring points for the same radius, treating them as miscellaneous will reduce overall face recognition rate. That's why one need to consider SNULBP's to increase face recognition rate.
- 4) From the above tables by looking at SNULBP column, it is clearly evident that face recognition rate is increasing gradually by increasing R. This is because as we increase R the LBP contains more number of NULBP's. Therefore one should consider the proposed SNULBP's for an accurate face recognition, as R increases.
- 5) There is no much variation for SNULBP in case 2.

Table 8: Recognition rate by considering new test images (not part of training database) for AT&T ORL database: Initial case.

(P,R)	ULBP	PLBP	SNULBP	SNULBP U ULBP	SNULBP U PLBP	SNULBP U SPLBP
(8,1)	85.00	86.55	17.50	87.50	90.75	82.55
(8,2)	86.25	88.75	28.75	89.75	92.55	88.75
(8,3)	95.00	94.50	40.00	96.00	94.75	92.50
(8,4)	91.25	93.15	50.00	95.50	96.75	93.25
(16,1)	88.75	93.25	18.75	96.25	97.55	92.50
(16,2)	95	96.55	22.50	97.00	97.75	94.55
(16,3)	97.5	96.75	31.75	97.50	98.25	96.25
(16,4)	92.5	97.15	41.25	98.15	98.75	92.25
Average	91.41	93.33	31.31	94.71	95.89	91.58

Table 9: Face recognition rate for AT&T ORL data base for case 2 (The combination of new and training set as test images).

(P,R)	ULBP	PLBP	SNULBP	SNULBP U ULBP	SNULBP U PLBP	SNULBP U SPLBP
(8,1)	91.25	92.03	19.15	93.63	94.13	90.03
(8,2)	91.88	93.13	31.13	94.45	95.45	93.13
(8,3)	96.25	96.50	42.35	96.75	96.75	95.00
(8,4)	94.38	95.55	51.55	95.45	97.75	95.38
(16,1)	93.13	95.85	25.42	96.88	97.75	95.00
(16,2)	96.25	97.03	24.54	97.25	97.85	96.03
(16,3)	97.50	97.13	34.25	97.50	98.15	96.88
(16,4)	95.00	97.33	42.56	98.83	98.25	94.88
Average	94.45	95.57	33.87	96.34	97.01	94.54



For a better visual comparison all the existing and proposed derivations on LBP are plotted for three considered databases in fig. 2, 3, 4 and 5.

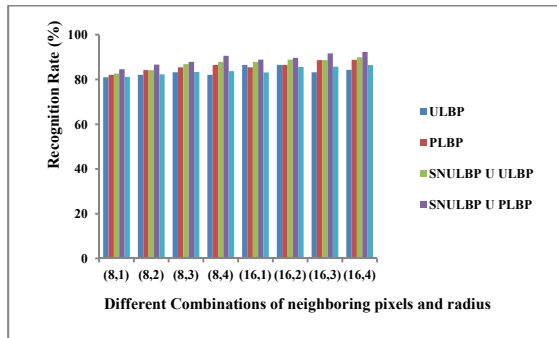


Figure 2. Face recognition rate for Yale database.

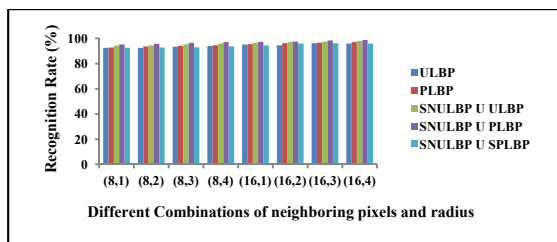


Figure 3. Face recognition rate for Indian database

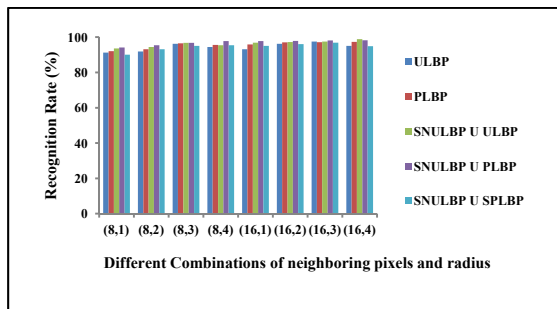


Figure 4. Face recognition rate for AT&T ORL database

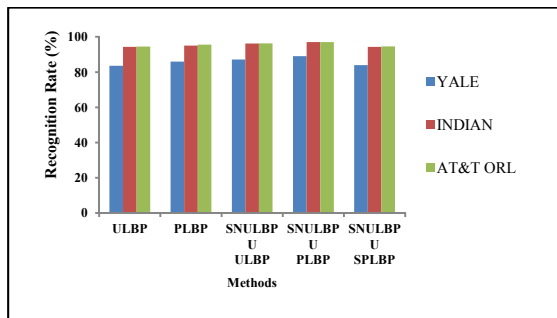


Figure 5. Face recognition rate comparison for Yale, Indian and AT&T ORL databases.

5. Conclusions

The present paper observed, the strong reason for considering NULBP's, because as we increase the P or R or both the number of NULBP's increases

abnormally. If one treats such a huge number of patterns as miscellaneous then definitely some image content will be lost and this degrades the overall performance. To overcome this and to deal with dimensionality the present paper derived SNULBP's. The proposed SNULBP's are stable, because it considered the transitions from two or more consecutive zeros or two or more consecutive one's. The graphs from the figures 2, 3 and 4 clearly indicates the proposed SNULBP U ULBP, SNULBP U PLBP has shown high performance when compared to ULBP alone for all considered databases. This clearly indicates the significance of the proposed SNULBP in improving overall face recognition rate. Further the SPLBP U SNULBP shown almost similar face recognition rate when compared to ULBP U SNULBP and ULBP alone.

6. Acknowledgments

The authors would like to express their gratitude to Dr. V. Vijaya Kumar, Director, Centre for Advanced Computational Research (CACR) & Dean, Dept. of CSE & IT at Anurag Group of Institutions, (AGOI) (Autonomous), Hyderabad, T.S., India and Dr. P. Rajeshwara Reddy, Chairman, Anurag Group of Institutions (AGOI), Hyderabad for providing necessary infrastructure for CACR at AGOI, which is bringing various research scholars across the nation to work under one roof. The CACR is providing a research platform for exchanging and discussing various views on different research topics. Authors extended their gratitude to Management of BVRIT Hyderabad College of Engineering for Women, Hyderabad, T.S. India, for promoting the young staff members towards research activities. Authors would like to thank authorities of Yale database, Indian database and AT&T ORL database.

7. References

- [1] Ahonen, Timo, Abdenour Hadid, and Matti Pietikainen. "Face description with local binary patterns: Application to face recognition." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 28, no. 12 (2006): 2037-2041.
- [2] Albiol, Alberto, David Monzo, Antoine Martin, Jorge Sastre, and Antonio Albiol. "Face recognition using HOG-EBGM." *Pattern Recognition Letters* 29, no. 10 (2008): 1537-1543.
- [3] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." In *Computer vision—ECCV 2006*, pp. 404-417. Springer Berlin Heidelberg, 2006.
- [4] Belhumeur, Peter N., Joao P. Hespanha, and David Kriegman. "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 19, no. 7 (1997): 711-720.



- [5] Bicego, Manuele, Andrea Lagorio, Enrico Grosso, and Massimo Tistarelli. "On the use of SIFT features for face authentication." In *Computer Vision and Pattern Recognition Workshop*, 2006. CVPRW'06. Conference on, pp. 35-35. IEEE, 2006.
- [6] Cao, Zhimin, Qi Yin, Xiaoou Tang, and Jian Sun. "Face recognition with learning-based descriptor." In *Computer Vision and Pattern Recognition (CVPR)*, 2010 IEEE Conference on, pp. 2707-2714. IEEE, 2010.
- [7] Cox, David, and Nicolas Pinto. "Beyond simple features: A large-scale feature search approach to unconstrained face recognition." In *Automatic Face & Gesture Recognition and Workshops (FG 2011)*, 2011 IEEE International Conference on, pp. 8-15. IEEE, 2011.
- [8] Dreuw, Philippe, Pascal Steingrube, Harald Hanselmann, Hermann Ney, and G. Aachen. "SURF-Face: Face Recognition Under Viewpoint Consistency Constraints." In *BMVC*, pp. 1-11. 2009.
- [9] Fathi, Abdolhossein, and Ahmad Reza Naghsh-Nilchi. "Noise tolerant local binary pattern operator for efficient texture analysis." *Pattern Recognition Letters* 33, no. 9 (2012): 1093-1100.
- [10] Heikkila, Marko, Matti Pietikainen, and Cordelia Schmid. "Description of interest regions with local binary patterns." *Pattern recognition* 42, no. 3 (2009): 425-436.
- [11] Heisele, Bernd, Purdy Ho, Jane Wu, and Tomaso Poggio. "Face recognition: component-based versus global approaches." *Computer vision and image understanding* 91, no. 1 (2003): 6-21.
- [12] Jain, Vidit, and Amitabha Mukherjee. "The Indian face database." URL [http://vis-www.cs.umass.edu/\\$ sim \\$ vidit/{I} ndian {F} ace {D} atabase\(2002\)](http://vis-www.cs.umass.edu/$ sim $ vidit/{I} ndian {F} ace {D} atabase(2002)).
- [13] Liao, Shu, Max WK Law, and Albert CS Chung. "Dominant local binary patterns for texture classification." *Image Processing, IEEE Transactions on* 18, no. 5 (2009): 1107-1118.
- [14] Liu, Li, Lingjun Zhao, Yunli Long, Gangyao Kuang, and Paul Fieguth. "Extended local binary patterns for texture classification." *Image and Vision Computing* 30, no. 2 (2012): 86-99.
- [15] Lowe, David G. "Distinctive image features from scale-invariant keypoints." *International journal of computer vision* 60, no. 2 (2004): 91-110.
- [16] M. Chandra Mohan, V. Vijaya Kumar, "A New Method of Face Recognition Based on Texture Feature Extraction on Individual Components of Face" *International Journal of Signal and Image Processing (IJSIP)*, Vol. 1-2010/Iss.2, pp. 69-74, March, 2010.
- [17] M. Chandra Mohan, V. Vijaya Kumar, B. Sujatha, "Classification of Child and Adult Based on Geometric Features of Face Using Linear Wavelets" *International Journal of Signal and Image Processing (IJSIP)*, Vol.1-2010/Iss.3, 2010.
- [18] M. Chandra Mohan, V. Vijaya Kumar, U.S.N. Raju, "New Face Recognition Method Based On Texture Features Using Linear Wavelet Transforms", *IJCSNS International Journal of Computer Science and Network Security*, VOL.9 No.12, December 2009.
- [19] Maturana, Daniel, Domingo Mery, and Alvaro Soto. "Face recognition with local binary patterns, spatial pyramid histograms and naive Bayes nearest neighbor classification." In *Chilean Computer Science Society (SCCC)*, 2009 International Conference of the, pp. 125-132. IEEE, 2009.
- [20] Meyers, Ethan, and Lior Wolf. "Using biologically inspired features for face processing." *International Journal of Computer Vision* 76, no. 1 (2008): 93-104.
- [21] Moghaddam, Baback, and Alex Pentland. "Probabilistic visual learning for object representation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 19, no. 7 (1997): 696-710.
- [22] Murty, Gorti Satyanarayana, J. Sasi Kiran, and V. Vijaya Kumar. "Facial Expression Recognition Based on Features Derived From the Distinct LBP and GLCM." *International Journal of Image, Graphics and Signal Processing (IJIGSP)* 6, no. 2 (2014): 68.
- [23] Nanni, Loris, Sheryl Brahmam, and Alessandra Lumini. "A simple method for improving local binary patterns by considering non-uniform patterns." *Pattern Recognition* 45, no. 10 (2012): 3844-3852.
- [24] Nguyen, Hieu V., and Li Bai. "Cosine similarity metric learning for face verification." In *Computer Vision-ACCV 2010*, pp. 709-720. Springer Berlin Heidelberg, 2011.
- [25] Nguyen, Hieu V., Li Bai, and Linlin Shen. "Local gabor binary pattern whitened pca: A novel approach for face recognition from single image per person." In *Advances in Biometrics*, pp. 269-278. Springer Berlin Heidelberg, 2009.
- [26] Ojala, Timo, Matti Pietikainen, and David Harwood. "A comparative study of texture measures with classification based on featured distributions." *Pattern recognition* 29, no. 1 (1996): 51-59.
- [27] Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Gray scale and rotation invariant texture classification with local binary patterns." In *Computer Vision-ECCV 2000*, pp. 404-420. Springer Berlin Heidelberg, 2000.
- [28] Oliver, Arnau, Xavier Llado, Jordi Freixenet, and Joan Marti. "False positive reduction in mammographic mass detection using local binary patterns." In *Medical Image Computing*



- and Computer-Assisted Intervention–MICCAI 2007, pp. 286-293. Springer Berlin Heidelberg, 2007.
- [29] Perlibakas, Vytautas. "Distance measures for PCA-based face recognition." *Pattern Recognition Letters* 25, no. 6 (2004): 711-724.
- [30] Reddy, P., B. Eswara Reddy, and V. Vijaya Kumar. "Fuzzy Based Image Dimensionality Reduction Using Shape Primitives For Efficient Face Recognition." *ICTACT JOURNAL ON IMAGE AND VIDEO PROCESSING*, NOVEMBER 2013, VOLUME: 04, ISSUE: 02, 695-701
- [31] Rodriguez, Yann, and Sebastien Marcel. "Face authentication using adapted local binary pattern histograms." In *Computer Vision–ECCV 2006*, pp. 321-332. Springer Berlin Heidelberg, 2006.
- [32] Ruiz-del-Solar, Javier, Rodrigo Verschae, and Mauricio Correa. "Recognition of faces in unconstrained environments: a comparative study." *EURASIP Journal on Advances in Signal Processing* 2009 (2009): pp. 1–20
- [33] Samaria, Ferdinando S., and Andy C. Harter. "Parameterisation of a stochastic model for human face identification." In *Applications of Computer Vision, 1994.*, Proceedings of the Second IEEE Workshop on, pp. 138-142. IEEE, 1994.
- [34] Shan, Caifeng, Shaogang Gong, and Peter W. McOwan. "Facial expression recognition based on local binary patterns: A comprehensive study." *Image and Vision Computing* 27, no. 6 (2009): 803-816.
- [35] Tan, Xiaoyang, and Bill Triggs. "Enhanced local texture feature sets for face recognition under difficult lighting conditions." *Image Processing, IEEE Transactions on* 19, no. 6 (2010): 1635-1650.
- [36] Turk, Matthew, and Alex Pentland. "Eigenfaces for recognition." *Journal of cognitive neuroscience* 3, no. 1 (1991): 71-86.
- [37] V.Vijaya Kumar, K.Srinivasa Reddy, V.V.Krishna, "Face recognition using prominent LBP model", *International Journal of Applied Engineering Research (IJAER)*, Volume 10, No.2(2015), pp 4373-4384.
- [38] Vu, Ngoc-Son, and Alice Caplier. "Enhanced patterns of oriented edge magnitudes for face recognition and image matching." *Image Processing, IEEE Transactions on* 21, no. 3 (2012): 1352-1365.
- [39] Wolf, Lior, Tal Hassner, and Yaniv Taigman. "Descriptor based methods in the wild." In *Workshop on Faces in'Real-Life'Images: Detection, Alignment, and Recognition*. 2008.
- [40] Wolf, Lior, Tal Hassner, and Yaniv Taigman. "Similarity scores based on background samples." In *Computer Vision–ACCV 2009*, pp. 88-97. Springer Berlin Heidelberg, 2010.
- [41] Wright, John, and Gang Hua. "Implicit elastic matching with random projections for pose-variant face recognition." In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 1502-1509. IEEE, 2009.
- [42] Yale Database, <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>
- [43] Zhang, Wenchao, Shiguang Shan, Wen Gao, Xilin Chen, and Hongming Zhang. "Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition." In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 1, pp. 786-791. IEEE, 2005.
- [44] Zhao, Guoying, and Matti Pietikainen. "Dynamic texture recognition using local binary patterns with an application to facial expressions." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 29, no. 6 (2007): 915-928.
- [45] Zhao, Wenyi, Rama Chellappa, P. Jonathon Phillips, and Azriel Rosenfeld. "Face recognition: A literature survey." *Acm Computing Surveys (CSUR)* 35, no. 4 (2003): 399-458.
- [46] Zhou, Hui, Runsheng Wang, and Cheng Wang. "A novel extended local-binary-pattern operator for texture analysis." *Information Sciences* 178, no. 22 (2008): 4314-4325.
- [47] Zou, Jie, Qiang Ji, and George Nagy. "A comparative study of local matching approach for face recognition." *Image Processing, IEEE Transactions on* 16, no. 10 (2007): 2617-2628.



Biographies



K. Srinivasa Reddy received Masters Degree in M.Tech. from SRM University, Chennai, T.N., India and pursuing Ph.D from JNTUA, Anantapuram, A.P., India, in Computer Science & Engineering under the guidance of Dr. V. V. Krishna. At present he is doing his research at Centre for Advanced Computational Research (CACR) of Anurag Group of Institutions (AGOI), Hyderabad, T.S, India. He is working as Associate Professor in BVRIT Hyderabad College of Engineering For Women, Hyderabad, T.S, India. His research interests include Image Processing, facial image and texture analysis. He has published research papers in various National, International conferences and Journals. He is a life member of ISTE, CSI.

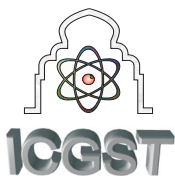


Dr V. Venkata Krishna received the B.Tech. (ECE) degree from Sri Venkateswara University. He completed his M. Tech. (Computer Science) from JNT University. He received his Ph.D. degree in Computer Science from Jawaharlal Nehru Technological University (JNTU), Hyderabad, India in 2004. He worked as Professor and Head of the Department for 10 years in Mahatma Gandhi Institute of Technology, Hyderabad. He worked as Principal for Vidya Vikas College of Engineering, JNTU, Hyderabad and Chaitanya Institute of Science & Technology, JNTU, Kakinada. At present He is working as Principal in Godavari Institute of Engineering and Technology, JNTUK, Rajahmundry, A.P., India. He is an advisory member for many Engineering colleges. He Published more than 30 research publications in various National, International journals and conferences. Presently he is guiding 10 research scholars. He is a life member of ISTE and CSI. He is a member of Srinivasa Ramanujan Research Forum-GIET.



Dr. B. Eswara Reddy graduated in B.Tech (Computer Science and Engineering) from Sri Krishna Devaraya University in 1995. He received Masters Degree in M.Tech. (Software Engineering) from JNT University, Hyderabad, in 1999. He received Ph.D in Computer Science & Engineering from JNT University, Hyderabad, in 2008. He has guided 3 research scholars for Ph.D. He served as Assistant Professor from 1996 to 2006 and as Associate Professor from 2006 to 2012. He is working as Professor of CSE Dept., at JNTUA College of Engineering, Anantapuram since 2012 and currently acting as coordinator for Master of Science in Information Technology (MSIT) programme offered at JNTU Anantapuram, currently he is Vice-Principal, J.N.T.U.A. College of Engineering, Anantapuram. He has more than 50 Publications in various International Journals and Conferences. He is one of the author's of the textbooks titled 'Programming with Java' published by Pearson/Sanguine Publishers and 'Data Mining' published by Elsevier India. His research interests include Pattern Recognition & Image Analysis, Data Mining and Cloud Computing. He is a life member of ISTE, ISCA, CSI, Fellow IE (India) and IEEE.





WATERMARKING SCHEMES FOR COPYRIGHT PROTECTION OF GEOSPATIAL DATA, A REVIEW

Sangita Zope-Chaudhari¹, Parvatham Venkatachalam², Krishna Mohan Buddhiraju³

¹Research Scholar, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India.

Email: sangita.z@iitb.ac.in

²Professor, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India.

Email: pvenk@csre.iitb.ac.in

³Professor, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India.

Email: bkmohan@csre.iitb.ac.in

Abstract

Rapid developments in information technology have promoted the extensive use of geospatial data in many fields, ranging from defense and military to personalized applications such as car navigation systems. However, with extensive use, the security of such data becomes even more vital and important. Digital watermarking has been used for the copyright protection of geospatial data as well as for its authentication and origin tracing. However, to ensure complete protection, it is important to ensure compatibility between the watermarking scheme and the features of different geospatial data. In this paper, various digital watermarking algorithms used for the copyright protection of geospatial data are reviewed, and further improvements are identified.

Keywords: Copyright protection, watermarking, geospatial data, robustness.

1. Introduction

In last two decades, the growth of high-speed computer networks and the World Wide Web have promoted many scientific and social breakthroughs, including electronic publishing, real-time information delivery, data sharing, collaboration among computers, and digital repositories. Today, computers have enabled efficient storage and manipulation of high-quality digital maps. However, the distribution of digital maps via the Internet leads to unlimited copying, which in turn poses a great threat to the map owners, thereby making it imperative to protect the digital copyrights of such data. One such method of copyright protection is digital watermarking.

Generally, geospatial data is modeled using raster and vector data models. In the raster data model, data is represented as a matrix consisting of uniformly sized cells, where each cell is identified by a unique row and column number. The cell contains a number or code, which represents the value type of the attribute being mapped. Raster models are similar to images; therefore, image watermarking algorithms can be directly used for the copyright protection of raster data. On the other hand, vector data is modeled using geometrical features such as points, polylines, and polygons. A topological relationship is employed to demonstrate the association between all these features. Geospatial data is different from other digital data because of its spatial characteristics. Thus, the requirements for the watermarking of geospatial data are also different. For example, in the case of vector data watermarking, it is necessary to preserve the precision and topological relationship of the points; maintain positional accuracy; provide robustness against attacks; and ensure that the watermarking is invisible. The visual quality of a vector map should not be affected by the watermark. Also, the watermark should not lead to element deformation. Similarly, in the case of raster data, in addition to ensuring invisibility and robustness against attacks, watermarking algorithms must be near-lossless. That is, the distortion in the watermarked data should be below a prescribed threshold so that the resulting analysis, such as the classification based on watermarked data, will result in correct classes. Also, the watermarking technique should not distort specific areas in the geospatial raster data [1].



Depending on the type of geospatial data to be watermarked, digital watermarking algorithms are classified into raster (also referred as satellite data) and vector data algorithms. They are further classified into spatial and transform domain algorithms on the basis of the embedding method. Spatial domain algorithms directly alter coordinates (or pixels) to hide watermark data, while transform domain algorithms alter the frequency transform of data elements to embed watermark data. Further, transform domain includes discrete cosine transform (DCT), discrete Fourier transform (DFT), and discrete wavelet transform (DWT). In general, transform domain algorithms are more robust than spatial domain algorithms, albeit less accurate. This paper provides an overview of the digital watermarking algorithms for copyright protection of geospatial data. The existing algorithms are evaluated by considering various attacks specific to geospatial vector and raster data. Further, given that fidelity requirements for geospatial data are different from those for multimedia data, fidelity criteria such as peak signal-to-noise ratio (PSNR) and mean square error of the algorithms are also compared.

This paper is organized as follows: the next section describes distinctive features of geospatial data watermarking. Sections 3 and 4 analyse the existing algorithms for vector and raster data watermarking, respectively. Section 5 provides an overview of some research directions, and Section 6 presents the conclusions

2. Features of Geospatial Data Watermarking

Geospatial data watermarking differs from general multimedia watermarking because of the distinct characteristics of geospatial data: structure, fidelity, robustness against attacks, and a spatial relationship between data elements. Although image and raster data have a common data structure (grid of pixels), they differ in terms of fidelity and attacks.

2.1. Data Structure

Vector data has two components: spatial data and attribute data. Spatial data describes the location of geographical objects, which consists of basic types such as points, polylines, and polygons. Each type of spatial data is often organized in a hierarchy of layers. On the other hand, attribute data represents the properties of geographical objects. A watermark is embedded in the coordinates of spatial data since attribute data usually does not allow manipulation. As vector data has a floating-point coordinate sequence, it is difficult to directly apply classical multimedia watermarking methods to vector data. The distortion produced by watermark embedding should be in the allowable range of precision tolerance to avoid the degradation of vector data.

Watermarking for panchromatic raster images is similar to image watermarking. However, for multispectral satellite images, the watermark should be embedded in at

least three bands so as to build a composite image for analysis (classification).

2.2. Fidelity

Fidelity is the perceptual similarity between the watermarked image and the original image. In multimedia data watermarking, fidelity is measured using human perception or via indicators such as signal-to-noise ratio (SNR), mean square error, correlation, and just noticeable differences. However, for geospatial data, the aforementioned parameters are not adequate [2]. SNR only gives statistical values of changes in the entire data; it does not guarantee that the vertex error is within the precision tolerance of vector data. To evaluate the fidelity of vector data, shape distortion and a topological relationship between them should be considered. Spatial relations between the objects of a map such as adjacency, inclusion or exclusion, and intersection are described by a topological relationship. It is used for querying spatial data and rebuilding geographical objects. Redundant vertices and spatial features of vertices are modified during the watermark embedding process. This results in changes of topological relationships. Similarly, when evaluating the fidelity of satellite raster data, the classification accuracy should also be considered along with fidelity parameters for the watermarking of multimedia data.

2.3. Attacks

The robustness of a watermarking algorithm refers to its ability to oppose some common data processing operations, known as attacks. The nature of attacks relevant to geospatial data is different from that for general multimedia data. However, some of the multimedia data attacks can be directly applied on a raster dataset [3, 4].

2.3.1. Geometric Transformations

Geometric transformations include translation, scaling, rotation, and reflection. Such an attack is relatively easy to defend in vector data watermarking schemes as vector data has properties of non-loss scaling and rotation. In the case of raster data, geometric transformations do not result in the loss of resolution but may introduce a smoothing or ringing effect.

2.3.2. Editing Attacks

In vector data, this type of an attack involves editing operations performed on vertices or coordinates. It includes the addition of new vertices (interpolation), removal of vertices (cropping and simplification), and reordering of vertices. Cropping and simplification attacks are very serious and crucial for geospatial data. For high-resolution satellite images, the cropping of even very small images results in substantial data loss.

2.3.3. Noise Distortion

A random, uncorrelated small value is added (or multiplied) in every pixel or coordinate of the watermarked data so that the changes are unnoticeable. In vector data, noise can be introduced by adding a random value (in centimeters) to coordinate values.



2.3.4. Compression

It simplifies the datasets by retaining the most significant features. Compression schemes such as Joint Photographic Experts Group (JPEG) and set partitioning in hierarchical trees are widely employed to evaluate the robustness of raster data watermarking methods. For vector data, some of the commonly used compression methods include the Douglas–Peucker algorithm, interval point selection, and vertical point deflection angle. A stepwise illustration of the Douglas–Peucker method is shown in figure 1.

Two extreme endpoints of a polyline are considered. An appropriate value for approximation tolerance ϵ is assumed. Point C is found on the polyline with the greatest distance b to line a (line joining extreme points). If b is lower than ϵ , then the polyline is approximated using the straight line a . Otherwise, the whole process is recursively repeated for each segment (the first endpoint is mapped to C, and C is mapped to the second endpoint) to obtain the approximation [5].

2.3.5. Enhancement Attacks

This type of attack is mainly used for raster datasets. It includes enhancement techniques such as low-pass filtering, sharpening, histogram modification, gamma correction, and restoration. The sharpening attack is very effective in some watermarking methods as it is extremely effective at detecting high-frequency noise introduced by watermarking methods.

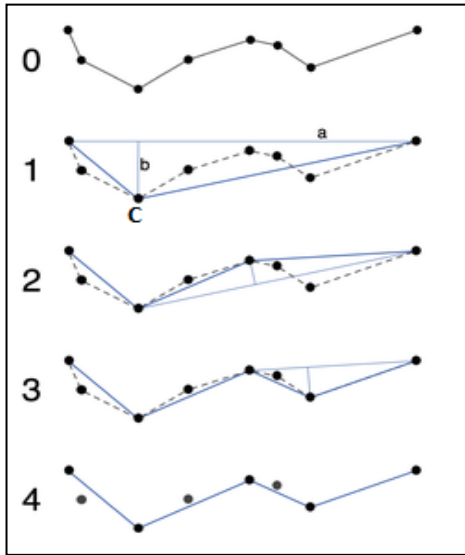


Figure 1. The Douglas–Peucker simplification process

2.3.6. Projection Transformation Attack

This type of transformation refers to the conversion of geospatial vector data from one projection to another projection and its conversion back to the original by inverse transformation.

2.3.7. Format Conversion

Whenever data is converted from one file format to another, some data is typically lost. However, whenever

data is converted from one format to Shapefile format, vector format (VEC), drawing exchange format and vice versa, marginal data loss is observed, which does not affect extracted results. On the other hand, for text format conversion, more data is probably lost, and it becomes difficult to extract the embedded watermark. In case of a raster dataset, format conversion can be done using JPEG, Tagged Image File Format (TIFF), Graphics Interchange Format (GIF) and Arcinfo GRID file formats.

3. Vector Data Watermarking

In the field of geographic information system (GIS), most of the research studies are focused on digital watermarking for copyright protection, which enables the secure distribution of geospatial data. This section provides an overview of vector data watermarking algorithms for spatial and transform domains.

3.1. Spatial Domain Algorithm

Vector data can be modeled using features such as points, polylines, and polygons. A topological relationship is employed to demonstrate the association between these features. Spatial-domain-based vector data watermarking approaches use these features and topological relations to embed a watermark. In the case of spatial domain, simple watermarks can be embedded in the vector data by modifying the values of X–Y coordinates.

Ohbuchi et al. [6] have presented a correlation-detection-based watermarking method for the watermarking of 2D vector maps. They have employed a uniform quadtree and modified quadtree division to partition the input vector map into rectangles containing a certain amount of vertices. Embedding information is repeated multiple times to form a watermark of a length similar to the number of divided rectangles in the map. Watermark bits are represented as bipolar values. In each rectangle, only a single bit of the watermark is embedded. For a particular rectangle, data embedding is performed using the additive approach.

$$C_i'' = C_i + b_i * PS_i * \alpha \quad (1)$$

Here, C_i and C_i'' are the original and watermarked coordinates, respectively. b_i is the i^{th} bit of the watermark ($b_i \in \{-1, 1\}$), PS_i is the i^{th} bit of the pseudo-random sequence ($PS_i \in \{-1, 1\}$), and α is the amplitude modulation (embedding strength). In the decoding procedure, the original map and the same pseudo-random sequence are required. The watermark is extracted using a correlation between the original and watermarked vertices. Ohbuchi et al. [6] have considered a value of α as 1 for all test cases. This will always lead to the displacement of vertices in either direction by one, resulting in visual distortion in the vector map. Moreover, this will surely violate the topological relationship. This algorithm provides good resilience against geometric attack, additive random noise (with amplitudes less than the embedding strength), vertex insertion or removal, and vertex reordering but less resilience against cropping attacks. The same method has been adopted by Wang et al. [7] to divide the vector map into different areas. All



points in the rectangles are converted to x and y direction components (V_x and V_y) of the vector set. They are calculated by determining the distance between the $(i+1)^{th}$ and i^{th} point in the rectangle. Then, the components are divided into even and odd intervals. Finally, depending on the value of watermarking bits, the watermark is embedded into x-y components. For large intervals, this scheme produces greater distortion and is less robust against noise attacks. To maximize watermark capacity, adjacent rectangles can be merged. This scheme provides acceptable distortion. Robustness is satisfactory for map simplification, interpolation, vertex disarrangement, geometric transformation, and noise attack. These attacks do not affect watermark extraction.

Some schemes use tolerance as a key parameter. The vector map is divided into a series of grids at a height of $4/3$ tolerance. In each grid, two lines are drawn at $1/3$ tolerance and 1 tolerance, and these two lines are marked as line 0 and line 1, respectively. Depending on the watermark bits, the watermark (metadata of vector data) is embedded by moving the vertices in the grids to line 0 or line 1. This scheme is only robust against polyline simplification, noise, and moving and cropping attacks [8]. Block codes are used to ensure the correctness of the extracted watermark information. A similar approach has been adopted in previous studies [9, 10].

In a vector map, there is a high correlation among the neighbouring vertices of the same feature. A polyline feature has been used for embedding the watermark by Cao et al. [11]. Highly correlated vertices of a polyline are grouped together, and their median is calculated. Watermark bits are iteratively embedded in that median value of each group. This scheme reduces the distortion caused by high correlation between the vertices and provides high payload capacity. However, the robustness of this scheme remains untested. A blind watermarking scheme using the polyline or polygon characteristics of maps has been reported in [12]. The length or perimeter is calculated for all polylines or polygons in a map. Considering the uniform step (key) of length or perimeter dynamic range, polylines or polygons are divided in sections or groups. Depending on the watermarking bit value, the mean value of each section is changed. The watermark is inserted multiple times. The dynamic range of the length or perimeter and the key used to divide them into groups are required to extract the watermark from the watermarked data. This scheme is robust against attacks such as geometric transformation, vertex reordering, vertex addition or deletion, cropping, and noise.

Even though objects are rotated and translated, the angle between them should remain unchanged. This feature has been used in [13]. An interior angle is calculated by applying the cosine rule on three consecutive vertices of an object. A user-defined random table is generated for values of 0–9, which serves as a key in watermark embedding. The integer part of the calculated angle is changed using this random table, and this value is referred as $C_A(x)$. The user's own value, i.e., watermark $W(x)$, is generated by adding ASCII values of all the

characters in a copyright string. The watermark angle is calculated as follows:

$$\text{Watermark angle} = W(x) - C_A(x) \% W(x) \quad (2)$$

The watermarked coordinates are calculated using the changed angle obtained in equation 1. Figure 2 shows the original coordinate $M(m_1, m_2)$, the watermarked coordinate $B(b_1, b_2)$ and the difference between the original angle and the watermarked angle θ . Equation 2 shows the watermarked coordinates.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} + \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} m_1 - l_1 \\ m_2 - l_2 \end{bmatrix} \quad (3)$$

Watermark extraction requires a random table and user's own value as the watermarking key. The extraction follows the same method as watermark embedding. Even though Kim [13] has claimed that the topology of the watermarked data is corrected using the intersection test, it cannot be used to correct all types of topological errors. A drawback of this method is that if the interior angles are changed, the watermark cannot be extracted. It is resilient against geometric transformation, noise addition, and vertex addition attacks but not against a vertex deletion attack.

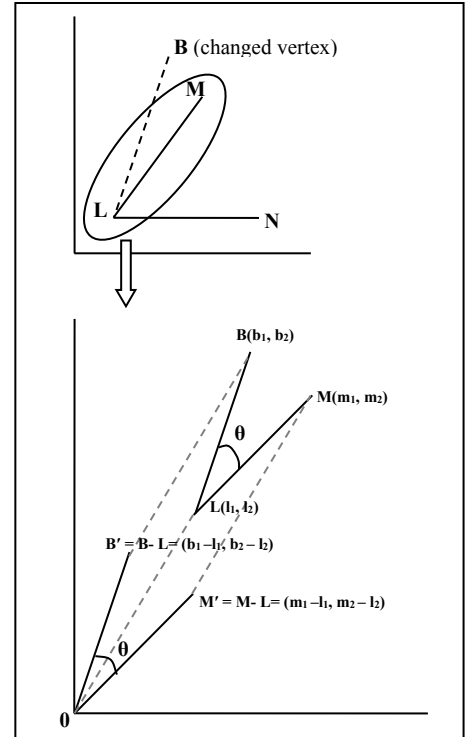


Figure 2. Calculation of watermarked coordinates

A lossless watermarking algorithm described in [14] is based on the global characteristics of a vector map. Vertices are categorized as feature (key) points, representing a polyline basic shape, and non-feature points, representing polyline details using the Douglas–Peucker algorithm [5]. A back-propagation neural



network is used to determine the watermarking parameter. Singular value decomposition (SVD) is utilized to calculate the watermarking parameter for non-feature points. In the watermark detection process, the detection keys are obtained through an exclusive OR (XOR) operation between lossless watermarking parameters and the copyright image. While this method performs satisfactorily against simplification and compression attacks, it provides good resilience against geometric attacks. The same concept has been used in [15] for developing a blind watermarking scheme for topographic maps. In the geospatial database, multiple layers are present, which indicate various themes. This approach selects three feature layers and key points from each layer. The selection of feature layers and key points is crucial, and they serve as key parameters for this method. The selection of key points can be performed using Voronoi diagrams, topological relations, and other geometrical measures. A string-based watermark is embedded into the coordinates of key points using the least significant bit algorithm. A similar procedure is performed for the extraction of the watermark from all feature layers to verify the correctness of the watermark. This scheme is robust against format exchange, vertex insertion or deletion, noise, and similarity transformation attacks; however, it has some limitations. First, the performance of the scheme depends on the selection of data layers. Second, it is not robust against data modification through map generalization.

Very few researchers have used topological relations for watermarking. Wang et al. [16] have designed a robust and blind watermarking algorithm that is based on the topological relations between map elements. In this approach, a minimum encasing rectangle (MER) for a map is calculated. With MER as a reference, the minimum bounding rectangle (MBR) of each polygon is calculated by the convex hull method. Each polygon is represented by the Hilbert code of its center. The message authentication code (MAC) is calculated by the hashing concatenation of the Hilbert code and a secret key, and all polygons are sorted on the basis of their MACs. Polygon pairs are selected in a way such that they do not overlap each other. A single watermark bit is inserted in every two polygons. A watermark is embedded in a spatial topological relation by slightly modifying the least significant bit (LSB) of the metric measures between each pair of polygons. Figure 3 shows flowchart of this scheme. Owing to the geometry invariance property of the spatial topological relation, this approach is robust against geometric, simplification, and interpolation attacks but not sufficiently robust against top cropping attacks. In a similar method, a zero watermark is generated by applying hashing on topological characteristics of the map data, and it is registered to an authority (third party) repository for notarization and authentication. Since it does not change the characteristics of the original data, this blind scheme is imperceptible and robust against projection transformation, coordinate transform, and data clipping [17].

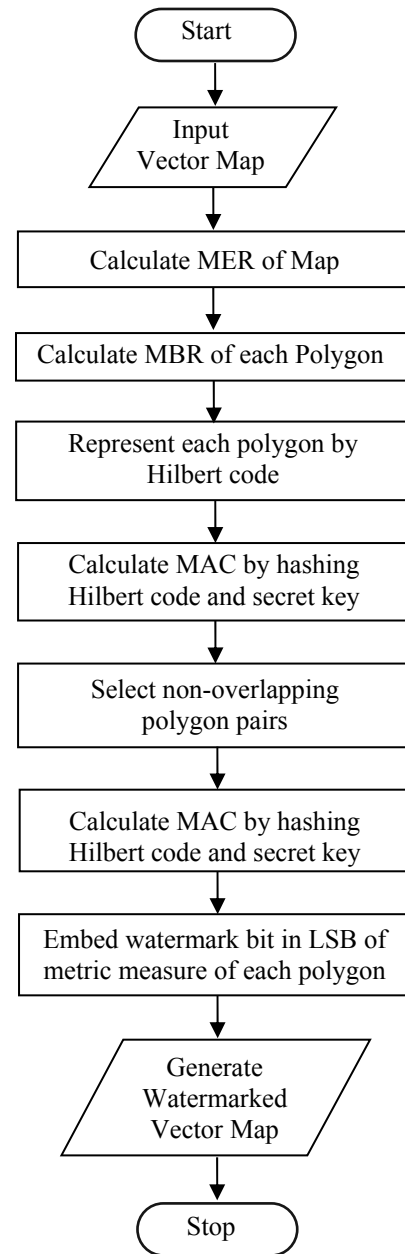


Figure 3. Flowchart for Wang et al. [16] technique

In addition to the above-mentioned characteristics, many other researchers have considered tolerance, error correction theory, and global characteristics of vector data for watermarking. A non-blind watermarking algorithm based on error tolerance has been proposed in [18]. The watermark survives and is successfully detected even if coordinates are changed within the tolerance range. The distortion can be easily controlled, and this scheme is robust against noise attack. An algorithm based on the data configuration of a vector map as well as the theory of error correction has been proposed in [19]. Features with more vertices are selected for watermark embedding. The vector angle is calculated for each vertex. The encoded watermark is embedded into a vertex depending on the vector angle and watermark bit value. This algorithm provides good resistance for attacks such as geometric transformation, vertex insertion or deletion, and local modifications.



However, it provides poor robustness against clipping attacks.

Marques et al. [20] have proposed a non-blind watermarking algorithm named RAWVec, in which the watermark (image) is inserted into the vector map coordinates by shifting the vertices. The detection of the watermark uses the original vector map to extract the watermark from the vector data. The original and extracted watermarks are compared using both a point pattern matching algorithm as well as visually, which increases the effectiveness of the system. Robustness to attacks is determined by the proposed point matching algorithm. It works well against noise addition, vertex reordering, and similarity transformation attacks; however, it is not resistant to projection transformation attacks. The method has been revised and renamed as SB-RAWVec, which is the semi-blind version of the earlier RAWVec algorithm. It uses the original map during detection but does not reveal it. However, the robustness of this scheme remains untested [21].

3.2. Transform Domain Algorithm

Similar to images, vector data can be also represented or stored in spatial and transform domains. To transfer vector data to its frequency representation, several reversible transforms such as DCT, DWT and DFT can be used. Watermarks can be embedded by modifying the transform domain coefficients of coordinates.

Let us first discuss DFT. There are a few algorithms that modify the magnitude and phase coefficients of DFT to embed watermarks. In the blind watermarking scheme proposed in [22], the DFT phase is quantized with step size and then the watermark is embedded in a quantized phase. The step size affects the visibility and robustness of map data. This method is not robust against scaling and vertex deletion attacks; however, it effectively resists other attacks such as translation, rotation, and format conversion. In a similar approach, a watermark is embedded into the DFT coefficient of the polygonal line vertex series. The detection of a watermark does not require the original polygonal line as this is a blind watermark detection method. Although this scheme is robust against geometric transformation, smoothing filtering, and noise attack, it is not robust against polygonal line cropping and vertex insertion or deletion [23]. In [24], DFT has been used for embedding a watermark in a set of polylines in a vector map. The steps of the scheme are listed below:

- (a) Each pair of N vertices is combined to get a complex number sequence :

$$C_k = \{x_k + iy_k \quad k = 0, \dots, N\}$$

DFT is applied on this complex data sequence.

- (b) A watermark of length m is considered as :

$$W_m \in \{0, 1\}.$$

- (c) The watermark is embedded using

$$A_l'' = |A_l| + S(-1)^{W_m} \quad (\lfloor \alpha N \rfloor \leq l \leq \lfloor \beta N \rfloor) \quad (4)$$

where A_l'' is the modified Fourier coefficient. A_l is the original Fourier coefficient, S is the embedding strength, α and β denote the lower and upper bound of the embedding frequency ($0 \leq \alpha < \beta \leq 1$), respectively. ($\lfloor \alpha N \rfloor$









- $\lfloor \beta N \rfloor$) represents the total embeddable bits. The watermarked vertex coordinates are calculated by taking the inverse DFT of the watermarked complex number.

- (d) The watermark is extracted by comparing Fourier coefficients and the original and watermarked data.

$$We = \begin{cases} 0 & (|A_l'| - |A_l| \geq 0) \\ 1 & (|A_l'| - |A_l| < 0) \end{cases} \quad (\lfloor \alpha N \rfloor \leq l \leq \lfloor \beta N \rfloor) \quad (5)$$

We evaluated this scheme and observed that the embedding strength is the key factor, and it should be carefully selected as the visual distortion of vector data is proportional to the embedding strength. We evaluated the robustness of this scheme against noise, compression, geometrical attacks (translation, scaling), format exchange, and coordinate addition or deletion. This algorithm exhibited good robustness against all these attacks (Table 1) but resulted in some inadequacies for cropping attack. Figure 4 shows the watermarked contour data. Figure 5 shows visual degradation as a function of embedding strength. Watermarked data was superimposed on the original data to indicate visual degradation. Red and blue lines show original and watermarked vector data, respectively.

Table 1 Robustness Evaluation of a Fourier-based Scheme (Kitamura et al., 2001)

Attacks		Extracted Watermark	Normalized Correlation (NC)
Noise			0.875840
Compression			0.866890
Translation			1
Scaling	Upscaling		1
	Down-scaling		1
Format Exchange			0.995714
Coordinate Addition			1
Coordinate Deletion (5%)			0.936790

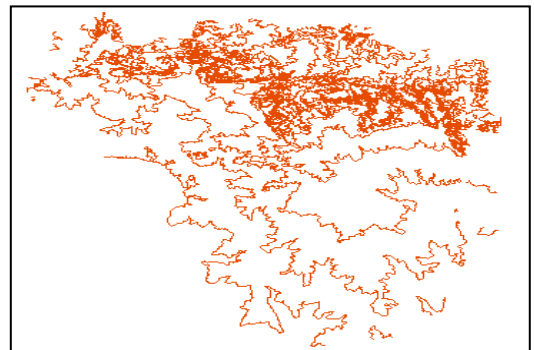


Figure 4. Watermarked contour data



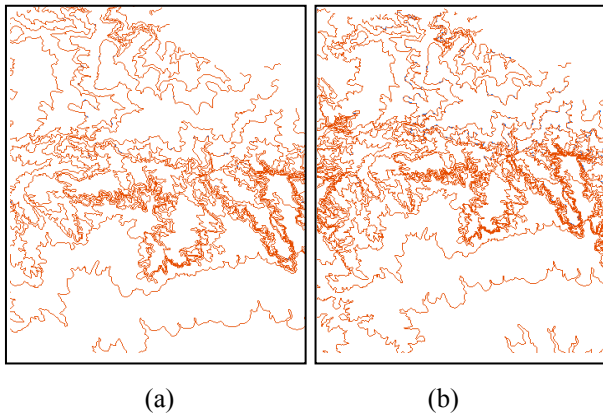


Figure 5. Visual degradation of watermarked contour data (Sectional enlarged view of the original and watermarked data) (a) for $p = 300$, $\alpha = 0.1$ and $\beta = 0.3$ (b) for $p = 500$, $\alpha = 0.1$ and $\beta = 0.3$

In the non-blind watermarking method proposed in [25]. Contour lines are used to embed a watermark. DFT is applied on the vertex coordinates, and watermark bits are embedded in the DFT phase sequence. The extracted watermark is compared with the original one by the autocorrelation method. No watermark is embedded in the first coefficient and the first DFT phase sequence as they get affected by translation and rotation. As the DFT phase is invariant to geometric transformations, the watermark can survive and can be detected under the geometric attack scenario. This method provides good imperceptibility and robustness against translation, scaling, vertex deletion, and format changing attacks. However, even a small change in the phase can introduce significant distortion in the vector data.

A DCT-based reversible and blind watermarking scheme has been proposed in [26]. It utilizes the high correlation property of vertex coordinates from the same feature. A watermark bit is embedded in a group of eight point DCT coefficient vertices. This algorithm not only provides high-capacity watermarking but also reduces map distortion. Wang et al. [27] have adopted the same scheme as that proposed in [26]. They used a threshold-based error estimation method to control the distortion of watermarked data. The watermark can be fully recovered under operations such as translation, scaling, vertex insertion or deletion, noise addition, cropping, and object scrambling. However, the algorithm still lacks high capacity and robustness against rotation attacks. The scheme described in [28] uses the feature domain along with the transform domain. In this scheme, MBR is computed and divided into uniform grids. The watermark is scrambled using Arnold permutation. Watermarking bits are embedded in medium-frequency DCT coefficients of a grid-weighted array for each grid. Map distribution is modified by taking inverse DCT. The same algorithm is applied for watermark extraction. This algorithm gives good imperceptibility and is also robust against geometric transformation and simplification attack. In a similar approach, feature vertices are extracted using the Douglas–Peucker algorithm to form feature images. The DCT transform is performed on this feature image, and a watermark is embedded into the middle- and low-frequency coefficients. Finally, the

inverse DCT transform is applied on the adjusted coefficients to obtain watermark data. This scheme provides good imperceptibility and is robust against some common attacks [29].

In last decade, many researchers have focused on wavelet transform in the field of watermarking, because it can better reflect the local features while being insensitive to local modification. Also, the window of the wavelet function can be changed by a scale factor. Li and Xu [30] have developed a wavelet-based blind watermarking scheme for the protection of vector maps. The magnitude and phase are calculated from detail coefficients at a particular decomposition level, and a character sequence is generated from the magnitude. Watermark bits are embedded into the character sequence. This algorithm is sufficiently robust to deal with attacks such as geometric transformation and noise addition. In a non-blind algorithm discussed in [31], watermark bits are embedded in low-frequency coefficients by applying integer wavelet transform on vector data. Flowchart of this scheme is shown in figure 6. This algorithm can effectively resist attacks such as noise, data compression, point deletion, and format exchange. An approach proposed in [32] uses the Douglas–Peucker algorithm to classify vertices as feature and non-feature points. It uses the Haar wavelet coefficients of feature points to embed the watermark. Feature points are calculated for each sub-region generated by an area subdivision process. This method is robust against geometric transformation as well as addition, deletion, and cropping of points. Also, map distortion is well controlled. However, this algorithm does not provide good robustness against irregular cropping.

Table 2 presents a comparative summary of select vector data watermarking algorithms in spatial and transform domains on the robustness against attacks and fidelity criteria. Shortcomings and unresolved issues in previous studies have been clearly highlighted. Vector data can be modeled using points, lines, and polygons as basic features, and only a few existing algorithms [12, 16, 17] have considered these features in the watermarking embedding process.

4. Raster Data Watermarking

Raster data mostly include satellite images, digital elevation models, and aerial photos. Many studies have been conducted on image watermarking for copyright protection; however, very few methods have been developed for watermarking of remote sensing images. Some of the methods proposed for image watermarking can also be applied to remote sensing images if they fulfil the requirements of satellite image watermarking. In the following sections, the methods available for watermarking remote sensing images are reviewed.

4.1. Spatial Domain Algorithm

Chauhan et al. [33] have proposed a blind watermarking algorithm for the copyright protection of satellite images.



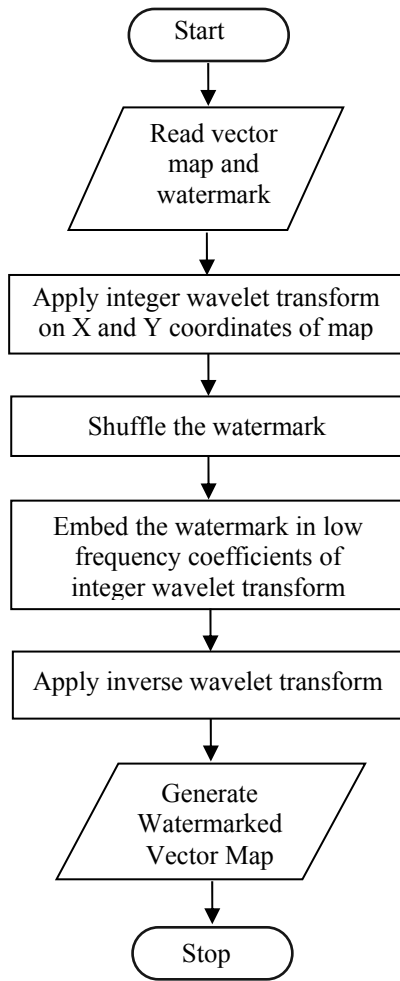


Figure 6. Flowchart for Zhu et al. [31] technique

A watermark is inserted without disturbing the vital areas of interest. Pixel values are quantized to 0 or 1 depending on the look-up table (LUT). A watermark is embedded in the pixels of the original image using mapping in the LUT. To find the corresponding position for embedding the watermark bit, a prime constant is used. In the extraction process, a watermark key, which is a combination of the LUT and prime constant, is used. As the spatial resolution of a satellite image is large, it is possible to embed multiple watermarks into it. Also, the LUT can be extended to deal with a tricolor or multicolor watermark using ternary or n-ary sequence instead of 0s and 1s in the LUT. The watermark is successfully extracted, and it gives good imperceptibility. However, the algorithm has not been evaluated against attack models. The watermarking tool “StegMark” was originally developed for robust and fragile multimedia image watermarking. Robustness is evaluated using cropping, rotation and compression. “StegMark Geo” is an invisible watermarking solution designed for the copyright protection of large-scale, high-resolution satellite images. It provides robustness against resizing as well as cropping and compression attacks [34].

A modified patchwork-based watermarking (MPW) scheme has been proposed in [35]. Based on spatial domain encryption, the MPW scheme is capable of

embedding a watermark with minimal manipulation of the original image pixel values. In this algorithm, n (watermark length) pairs of pixels (p_i, q_i) are selected satisfying two conditions: p_i and q_i should have the same intensity, and q_i should be from the neighbourhood $(3 \times 3, 5 \times 5)$ of p_i . p_i s and the corresponding q_i s are sorted. A watermark is converted into a binary one. The binary watermark and its complement are added in p_i and q_i , respectively, as shown below:

$$\begin{aligned} p'_i &= p_i + w_i \\ q'_i &= q_i + \bar{w}_i \end{aligned} \quad (6)$$

The key is formed using number of rows, number of pixels and relative distances of the watermarked pixels. Using distances in the given key, pixel values are retrieved in the extraction process. If the difference between p_i and q_i is +1, then the watermark is 1, and if the difference is -1, then the watermark has value 0. The MPW watermark embedding process does not create any visual artifacts and is imperceptible. The watermark retrieval process operates with the help of a key and does not require the original image. In this paper, the MPW-watermarked images are compared with DCT-based and wavelet-based watermarked images. The MPW watermark is imperceptible, and the embedding process does not introduce any visual artifacts. This method is robust against cropping and enhanced attacks but not against filtering and compression attacks. In the case of cropping, it is possible to retrieve the “partial” watermark even if the sub-image does not contain all the watermarked pixels.

Zhu et al. [36] have proposed an encryption-based watermarking scheme for the copyright protection of remote sensing images. A binary watermark is first scrambled using Arnold scrambling and then encrypted using a random number (key1). Blocks of the host image are encrypted using a random binary matrix (key2). Watermark bits are embedded in blocks depending on the values of the watermark bits. After embedding, decryption is carried out using key2 to get the watermarked image. Figure 7 shows the flowchart for watermark embedding procedure. The exact reverse procedure is used for the extraction of the watermark. This scheme provides good robustness for JPEG compression, noise and filtering; however, it does not perform well against cropping attacks.

A tree-structured vector quantizer-based semi-fragile watermarking approach has been employed in [37]. In preprocessing, compression and decompression are applied to satellite images to reduce the noise of the sensors and to remove spurious values. Watermark bits are embedded by extracting one or more least significant bits of pixels from selective bands of multi- and hyperspectral data. As this is a semi-fragile scheme, any modification such as compression and filtering (tampering) can be detected. This scheme is robust against compression as well as copy and replace attacks. In most of the additive watermarking techniques, the watermark is inserted using the same embedding strength, which could result in the degradation of data quality, thereby resulting in misclassifications.



Table 2 Comparative Summary of Watermarking Algorithms for Vector Data Based on Robustness and Fidelity

Paper	Domain	Robustness								Fidelity		
		NO	GT	VIR	VR	CR	MS	ST	FC	ERR	SD	TR
Ohbuchi et al. (2002)	Spatial	^a ✓	✓	✓	✓	^b X	^c -	-	-	Yes	No	No
Wang et al. (2009)	Spatial	✓	✓	✓	✓	-	✓	-	-	Yes	No	No
Schulz and Voigt (2004)	Spatial	✓	✓	-	-	✓	✓	-	-	Yes	No	No
Shujun et al. (2007)	Spatial	-	✓	✓	-	✓	-	-	-	Yes	No	No
Cao et al. (2010)	Spatial	-	-	-	-	-	-	-	-	Yes	No	No
Huo et al. (2012)	Spatial	✓	✓	✓	✓	✓	-	-	-	No	No	No
J. Kim (2010)	Spatial	✓	✓	✓	-	-	-	-	-	Yes	No	Yes
Men et al. (2010)	Spatial	-	✓	-	-	-	X	-	-	No	No	No
Yan et al. (2011)	Spatial	✓	-	✓	-	-	-	✓	✓	No	No	No
Wang et al. (2012)	Spatial	-	✓	-	-	X	✓	-	-	Yes	No	No
Li et al. (2008)	Spatial	-	✓	-	-	✓	-	-	-	No	No	No
Voigt and Busch (2002)	Spatial	-	✓	✓	-	X	-	-	-	Yes	No	No
Marques et al. (2007)	Spatial	✓	-	-	✓	-	-	✓	-	No	No	No
Tao et al. (2009)	DFT	-	✓	-	-	-	-	-	✓	Yes	No	No
Solachidis and Pita (2000)	DFT	✓	✓	X	-	X	-	-	-	No	No	No
Kitamura et al. (2001)	DFT	✓	-	✓	✓	X	-	-	-	Yes	No	No
Xu and Wang (2010)	DFT	-	✓	✓	-	-	-	-	✓	Yes	No	No
Voigt et al. (2004)	DCT	-	-	-	-	-	-	-	-	Yes	No	No
Wang et al. (2011)	DCT	✓	✓	✓	✓	✓	-	-	-	Yes	No	No
Liang et al. (2011)	DCT	-	✓	-	-	-	✓	-	-	No	No	No
Lianquan and Qihong (2007)	DCT	✓	✓	-	-	✓	-	-	-	No	No	No
Li and Xu (2003)	DWT	✓	✓	-	-	-	-	-	-	Yes	No	No
Zhu et al. (2008)	DWT	✓	-	✓	-	-	✓	-	✓	Yes	No	No
Zhang et al. (2010)	DWT	-	✓	✓	-	✓	-	-	-	No	No	No

^a✓ indicates a positive response for robustness.^bX indicates a negative response for robustness.^cAn entry “-” indicates an unchecked attack or a fidelity parameter.**List of abbreviations:**

No – Noise, GT - Geometrical attacks (rotation, translation, and scaling), VIR - Vertex insertion or removal, VR - Vertex reordering, CR - Cropping, MS - Map simplification, ST- Similarity transformation, FC - Format conversion, ERR - Error is controlled or not, SD - Shape distortion is considered or not, TR - Topological relationship is considered or not.



Barni et al. [38] have proposed a minimum-impact-on-classifier watermarking technique, which is a variation of the classical additive watermarking framework. Different embedding strengths are used in spectral channels. In the watermarking process itself, the classification (K-means) is conducted, and the optimized embedding strength is estimated to reduce classification errors.

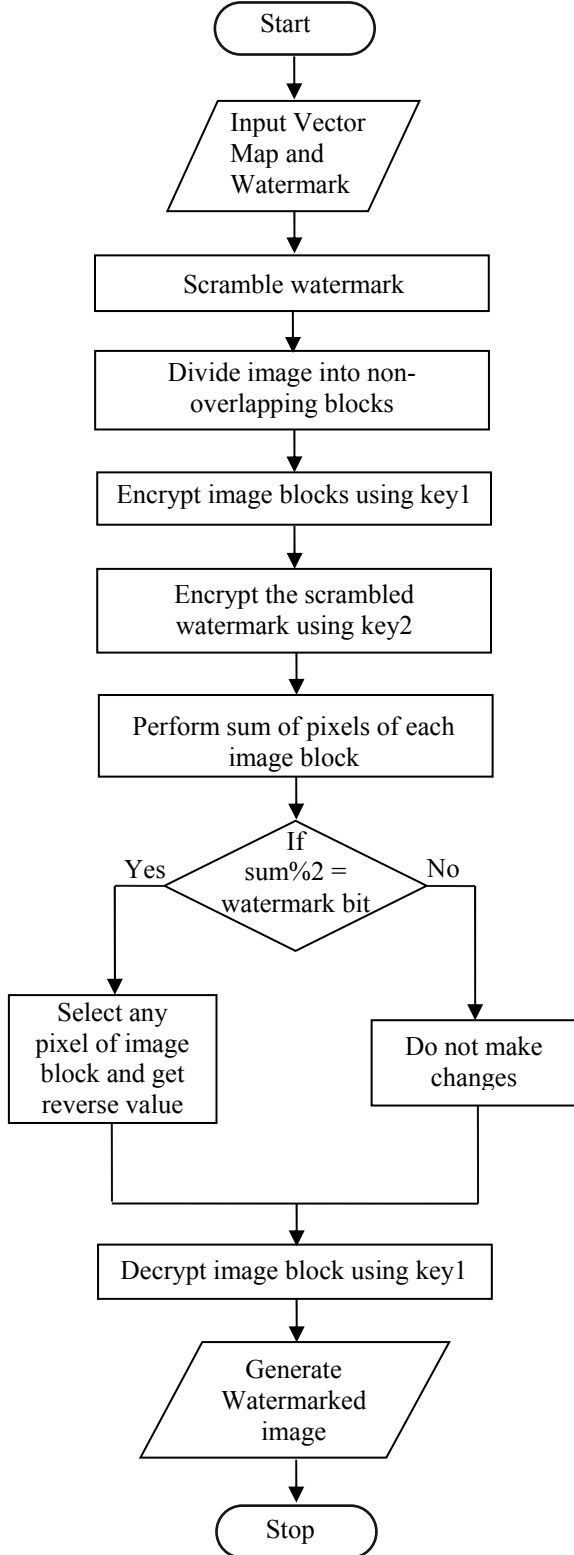


Figure 7. Flowchart for Zhu et al. [36] technique

4.2. Transform Domain Algorithm

Barni et al. [39] have proposed near-lossless watermarking for the copyright protection of remote sensing images. They have used the watermarking algorithms using DFT and DWT for still images proposed in (Barni et al., 2001b). The proposed solution rephrases the near-lossless concept by forcing the maximum absolute difference between the original and watermarked images. An error metric was used to measure the spectral distance and is defined as follows:

$$D_c^{m,n} = \max_i \left| x_i^{m,n} - \hat{x}_i^{m,n} \right| \quad (7)$$

Here, x is the original image, \hat{x} is watermarked image, (m, n) are the spatial coordinates of each pixel and i is its spectral value. In near-lossless watermarking, the spectral distance between the original and watermarked images is minimized, thereby causing less degradation in quality. This method has been evaluated to measure the impact of the cropping attack and classification. It was experimentally shown that robustness is less affected by cropping attacks. Robustness is better in a DWT-based scheme than in DFT. The classification results for DWT are better than those for DFT, and the overall classification results for both DWT and DFT using near-lossless watermarking are far better than other conventional watermarking techniques. In another study, Barni et al. [40] have proposed cartographic image watermarking using text-based normalization. In this watermarking process, the text orientation is first estimated by DFT. Second, the image is rotated so that the text assumes a given reference orientation. Third, the text is extracted from the image, and its size is estimated. Next, all the extracted characters are resized to the same size. Finally, the watermark is added in the magnitude of DFT coefficients in the medium portion of the frequency spectrum using the additive or multiplicative watermark embedding rule. This process is followed by inverse scaling and rotation to obtain the watermarked cartographic image. Li and Guang [41] have also used DFT to propose robust and blind watermarking using a segmented watermark. To resist common attacks, the invariant centroid of a normalized remote sensing image is determined. A square area is selected around the invariant centroid for watermark embedding. A pseudo-random sequence is used as a watermark, and it is segmented in two parts using a key. DFT is applied on the selected region. One segment of the watermark is embedded in the amplitude of DFT, while the other segment of the watermark is embedded into phases of DFT using an additive watermarking framework. Embedding strengths are determined in a way such that they balance invisibility and robustness. Finally, inverse DFT is applied to obtain the watermarked image. A similar method is used for the extraction of the watermark. This scheme is robust against compression, noise, and geometric transformation.

A semi-blind watermarking scheme for remote sensing images has been proposed in [42], where a wavelet transform domain was used to embed the watermark. This scheme retains special and spectral information of images. SVD is applied on any component of wavelet



decomposition (A, H, V or D). SVD matrices are required for watermark extraction. The Stirmark benchmark is used to check the robustness of the proposed algorithm. The algorithm is robust against attacks such as filtering, noise, geometric transformations, compression, and cropping. Compared with other wavelet-based algorithms, this algorithm works well for multimedia image watermarking. In another wavelet-based watermarking technique for satellite images described in [43], two zero watermarks are constructed from the image itself. One is constructed from a low-frequency component of the wavelet transform of a host image, while the other is constructed from the log-polar mapping image of the host image. This scheme is robust to attacks such as filtering, compression, noise, cropping and geometric transformation. In a remote sensing environment, analytic integrity is more important than perceptual data quality. A scheme using 3-level DWT decomposition has been shown in [44]. Here, only one sub-band at each level that has the maximum root means square (RMS) value is selected to embed the watermark. The watermark is embedded as indicated below:

$$bc'_i = bc_i (1 + \alpha w_i) \quad (8)$$

Here bc is the DWT coefficient, α is the watermarking strength, w is the watermark, and bc' is the watermarked DWT coefficient. Inverse DWT is applied to obtain watermarked raster data. The watermark is extracted with the help of the original raster data. This scheme is robust against cropping attacks. Also, the classification result is better as compared to other wavelet-based schemes as spatial features are preserved by high-value RMS bands. This scheme can be applied to multi- and hyper-spectral data.

Hsu and Chen [45] have proposed a feature-based robust watermarking scheme by utilizing scale-space feature points for watermark embedding as they are invariant to rotation, scaling, and translation. Scale invariant feature transform (SIFT) is used detect scale-space feature points. The circular area with a feature point as the center is used for watermark embedding. The watermark is added by applying a three-layer wavelet packet transform on the reorganized (rectangular) circular image. This scheme is robust to geometric transformations as well as compression, cropping, and enhancement operations. Even though this algorithm results in good classification accuracy, it does not survive under geometric corrections. Ho et al. [46] have adopted a fast Hadamard transform (FHT) in their character-based robust and invisible watermarking scheme. FHT is applied on pseudo-random-selected non-overlapping blocks of size 8×8 . Only 16 middle- and high-frequency AC components are selected for watermark embedding. Each character to be embedded is first represented by 7-bit ASCII and then in the Bose, Chaudhuri, and Hocquenghem code. Each watermarked bit (0 or 1) is converted into a stream of 16 numbers of $\{1, -1\}$ depending on its value and embedded into 16 AC coefficients. Figure 8 illustrates the flowchart of this approach. This algorithm is robust against cropping, noise, geometric transformations, filtering, and

compression. A similar algorithm was developed by Zhu and Ho [47] using the slant transform instead of FHT. The only difference was that instead of characters, they used an image as a watermark. Satellite images are highly textured. When a slant transform is applied to such images, it results into a good spread of middle-to-high frequencies with significant energies for robust watermarking. Watermark embedding has been carried out using the method proposed by Ho et al. [46]. This algorithm is only evaluated against JPEG compression attacks; however, its robustness against other attacks remains untested. As compared with the FHT-based algorithm, this algorithm gives good results for JPEG compression attack.

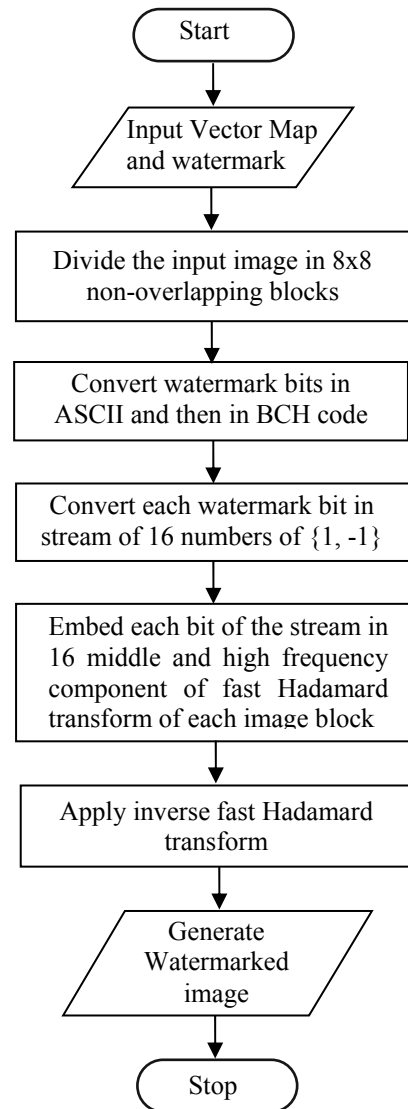


Figure 8. Flowchart for Ho et al. [46] technique

Digital elevation model (DEM) plays an important role in spatial data infrastructure. Not much effort has been taken for the copyright protection of DEM. Lui and Lv [48] have proposed DCT-based near lossless watermarking to protect DEM. The watermark is scrambled using the Hash algorithm and MD5 before embedding. The scrambled watermark is embedded into



8 x 8 medium-frequency DCT blocks using the additive approach. The terrain lines of DEM are selected for embedding the watermark. The strength of watermark embedding is a function of the slope (of DEM) and aspect (minimum and maximum errors), which will result in near-lossless watermarking. This scheme is robust against compression and cropping attacks.

Table 3 shows robustness against attacks and fidelity criteria of some of the raster data watermarking algorithms. Here as well, there is no standard set of attack models for evaluating watermarking algorithms. For fidelity evaluation, a well-known error measure, PSNR,

is used by almost all algorithms, but the classification error is considered by only a few approaches [39, 45, 35, 36, 44]. For the evaluation of raster data watermarking algorithms, all well-known attacks for image watermarking can be used. However, only few of them (compression and filtering) were used to check for the robustness of raster data watermarking algorithms. Thus, it is necessary to consider some more important attacks to evaluate the performance of raster data watermarking algorithms.

Table 3 Comparative Summary of Watermarking Algorithms for Raster Data Based on Robustness and Fidelity

Paper	Domain	Robustness							Fidelity	
		NO	GT	FL	CR	ENH	BR	COM	CERR	PSNR/ BER
Chauhan et al. (2002)	Spatial	-	-	-	-	-	-	-	No	No
Ho et al. (2001)	Spatial	-	✓	✓	✓	✓	-	✓	No	Yes
Kumari and Rallabandi (2008)	Spatial	-	-	X	✓	✓	-	X	Yes	No
Zhu et al. (2013)	Spatial	✓	-	✓	X	-	-	✓	No	Yes
Ruiz and Megias (2011)	Spatial	-	-	-	✓	-	-	✓	No	Yes
Barni et al. (2004)	Spatial	-	-	-	-	-	-	-	Yes	No
Barni et al. (2002)	DWT and DFT	-	-	-	✓	-	-	-	Yes	Yes
Barni et al. (2001c)	DFT	-	✓	-	✓	-	-	✓	No	No
Li and Guang (2012)	DFT	✓	✓	-	-	-	-	✓	No	Yes
Hemalatha et al. (2009)	DWT	✓	✓	✓	✓	-	-	✓	No	No
Jing et al. (2008)	DWT	✓	✓	✓	✓	-	-	✓	No	No
Zeigeler et al. (2003)	DWT	-	-	-	✓	-	-	-	Yes	No
Hsu and Chen (2012)	SIFT	-	✓	-	✓	✓	-	✓	Yes	Yes
Ho et al. (2003)	Hadamard	✓	✓	✓	✓	-	-	✓	No	Yes
Zhu and Ho (2003)	Slant	✓	✓	-	-	✓	-	✓	No	Yes
Liu and Lv (2008)	DCT	-	-	-	✓	-	-	✓	No	Yes

^a✓ indicates a positive response for robustness.

^bX indicates a negative response for robustness.

^c- An entry “-” indicates an unchecked attack or a fidelity parameter.

List of abbreviations:

No – noise, GT - geometrical attacks (rotation, translation, scaling and cropping), RS - resize, ENH - enhancement attack, BR -blurring, COM - compression, CERR - classification error is controlled or not, PSNR - peak signal-to-noise ratio, BER- bit error rate.



5. Research Directions

As observed in the literature, various algorithms are available for watermarking using spatial and frequency domains. However, each domain has certain limitations. Although the spatial domain retains accuracy, it provides weak robustness. On the other hand, the frequency domain provides good robustness but lacks imperceptibility. To increase robustness while preserving accuracy, it is essential to develop an integrated watermarking algorithm using different domains and/or a hybrid approach. Vector data can be modeled using points, lines, and polygons as basic features. Existing algorithms neither differentiate between points as well as linear and aerial features nor substantially focus towards the consideration of relations among geospatial objects.

At the distribution level, one of the most common attacks applied is cropping attack. In many watermarking algorithms, watermark bits are inserted in consecutive coordinates. When exploited by a cropping attack, such algorithms result in watermark destruction and hence unidentified copyright information. To deal with this issue, it is necessary to embed watermark bits into random vertex coordinates (or some key point coordinates) instead of consecutive vertex coordinates. Also, the same watermark can be inserted multiple times in vector data to make the algorithm more robust against regular and irregular cropping.

In previously specified literature, while evaluating the performance of a watermarking algorithm for vector data, the focus was particularly on only some common attacks such as geometrical transformation, noise, and vertex addition or deletion. However, no attempt was made to check whether the topological relationship was maintained, which is very important in vector data.

One of the important requirements of raster (remote sensing or satellite) data watermarking is that it should be near-lossless [49]. This requirement is only satisfied if the watermarked pixel values are extremely similar to the original pixel values (i.e., within some specified distance). If this requirement is not fulfilled, the percentage of misclassification of watermarked data is high. Also, only a few existing algorithms satisfy this constraint, and most of them use wavelet and Fourier transform coefficients for embedding the watermark.

Therefore, it is essential to develop algorithms that use other frequency domain methods while satisfying the above-mentioned constraint. For the evaluation of a raster data watermarking algorithm, it is necessary to consider some more important attacks other than compression and noise. Even if geospatial data is watermarked for copyright protection, intruders could insert their own watermark into already watermarked data and attempt to prove their ownership (copy attack). To deal with such situations, some cryptographic approaches such as identity-based encryption and digital signatures need to be used for solving authentication problems. In addition, fragile watermarking is the best solution for this purpose

6. Conclusion

GIS represents geographical information in terms of raster and vector forms. The compilation and management of geospatial data is expensive and time consuming. With the rapid development of the Internet and communication technology, it is easy to copy or distribute geospatial data. Therefore, copyright protection, authenticity, privacy, and spatial data source tracing have become important issues.

Based on the extensive review of available schemes, some unsolved issues such as robustness evaluation using various attack models have been observed, which deal with characteristics of geospatial data, distortion control and appropriate measures to evaluate the quality of geospatial data. The requirements for the watermarking of geospatial data is more stringent than those for the watermarking of multimedia data. Therefore, robust and invisible digital watermarking schemes need to be designed for fulfilling all specified requirements.

7. References

- [1] M. Barni, F. Bartolini, V. Cappellini, E. Magli, G. Olmo. Watermarking-based protection of remote sensing images: requirements and possible solutions. In proceedings of the International Symposium on Optical Science and Technology, 2001.
- [2] C. Lopez. Watermarking of digital geospatial datasets: a review of technical, legal and copyright issues. *International Journal of Geographical Information Science*, 16(6): 589–607, 2002.
- [3] L. Huang, W. Zhou, R. Jiang, A. Li. Data quality inspection of watermarked GIS vector map. In proceedings of the 18th International Conference on Geoinformatics, 2010.
- [4] Y. Zhou, A. Li, G. Lv. Research of robustness evaluation method for GIS vector data digital watermarking algorithm. In proceedings of the 18th International Conference on Geoinformatics, 2010.
- [5] D. Douglas and T. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Canadian Cartographer*, 10(2): 112–122, 1973.
- [6] R. Ohbuchi, H. Ueda, S. Endoh. Robust watermarking of vector digital maps. In proceedings of IEEE International Conference on Multimedia and Expo, 2002.
- [7] C. Wang, W. Wang, B. Wu, Q. Qin. A watermarking algorithm for vector data based on spatial domain. In proceedings of the 1st IEEE International Conference on Information Science and Engineering, 2009.
- [8] G. Schulz and M. Voigt. A high capacity watermarking system for digital maps. In proceedings of ACM Workshop on Multimedia and Security, 2004.
- [9] D. Shujun, L. Liang, C. Sen. Research on a digital watermarking algorithm suitable to vector map. In proceedings of IEEE International Conference on Automation and Logistics, 2007.
- [10] M. Voigt and C. Busch. Feature-based watermarking of 2D-vector data. In proceedings of SPIE



- International Conference on Security and Watermarking of Multimedia Content, 2003.
- [11] L. Cao, C. Men, X. Li. Iterative embedding-based reversible watermarking for 2D-vector maps. In proceedings of the 17th IEEE International Conference on Image Processing, 2010.
- [12] X. Huo, T. Seung, B. Jang, S. Lee, K. Kwon. A watermarking scheme using polyline and polygon characteristic of shapefile. In proceedings of the 3rd International Conference on Intelligent Networks and Intelligent Systems, 2010.
- [13] J. Kim. Robust vector digital watermarking using angles and a random table. *Advances in Information Sciences and Service Sciences*, 2(1): 79–90, 2010.
- [14] C. Men, L. Cao, X. Li, N. Wang. Global characteristic-based lossless watermarking for 2D-vector maps. In proceedings of the International Conference on Mechatronics and Automation, 2010.
- [15] H. Yan, J. Li, H. Wen. A key points-based blind watermarking approach for vector geo-spatial data. *Journal of Computers, Environment and Urban Systems*, 35(6): 485–492, 2011.
- [16] C. Wang, Z. Peng, Y. Peng, L. Yu, J. Wang, Q. Zjao. Watermarking geographical data on spatial topological relations. *Internal Journal of Multimedia Tools and Applications*, 57(1): 67–89, 2012.
- [17] A. Li, B. Lin, Y. Chen, G. Lo. Study on copyright authentication of GIS vector data based on zero watermarking. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, 37(4): 1783–1786, 2008.
- [18] M. Voigt and C. Busch. Watermarking 2D-vector data for geographical information system. In proceedings of SPIE International Conference on Security and Watermarking of Multimedia Contents, 2002.
- [19] F. Cheng, H. Yin, X. Zhang, D. Zhang. A digital watermarking algorithm for vector map. In proceedings of International conference on challenges in environmental science and computer engineering, 2010.
- [20] D. Marques, K. Magalhaes, R. Dahab. RAWVec-A method for watermarking vector maps. In proceedings of the 7th Brazilian Symposium on Information and Computer System Security, 2007.
- [21] K. Magalhaes, R. Dahab. SB-RAWVec-A semi-blind watermarking method for vector maps. In proceedings of IEEE International Conference on Communications, 2009.
- [22] S. Tao, X. Dehe, L. Chengming, S. Jianguo. Watermarking GIS data for digital map copyright protection. In proceedings of the 24th International Cartographic Conference, 2009.
- [23] V. Solachidis and I. Pitas. Watermarking polygonal lines using Fourier descriptors. *IEEE Computer Graphics and Applications*, 23(3): 44–51, 2000.
- [24] I. Kitamura, S. Kanai, T. Kishinami. Copyright protection of vector map using digital watermarking method based on discrete Fourier transform. In proceedings of IEEE International Symposium on Geosciences and Remote Sensing, 2001.
- [25] D. Xu, Q. Wang. The study of watermarking algorithm for vector geospatial data based on the phase of DFT. In proceedings of IEEE International Conference on Wireless Communications, Networking and Information Security, 2010.
- [26] M. Voigt, B. Yang, C. Busch. Reversible watermarking of 2D-vector data. In proceedings of ACM Multimedia and Security Workshop, 2004.
- [27] X. Wang, D. Huang, Z. Zhang. A DCT-based blind watermarking algorithm for vector digital maps. *Journal of Advanced Material Research*, 179(180): 1053–1058, 2011.
- [28] B. Liang, J. Rong, C. Wang. A vector maps watermarking algorithm based on DCT domain. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, 38(1): 118–121, 2011.
- [29] M. Lianquan and Y. Qihong. A digital map watermarking algorithm based on discrete cosine transform. *Journal of Computer Applications and Software*, Issue, 1(1): 146–148, 2007.
- [30] Y. Li and L. Xu. A blind watermarking of vector graphics images. In proceedings of 5th International Conference on Computational Intelligence and Multimedia Applications, 2003.
- [31] C. Zhu, C. Yang, Q. Wang. A watermarking algorithm for vector geo-spatial data based on integer wavelet transform. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, 37(B4): 15–18, 2008.
- [32] L. Zhang, D. Yan, S. Jiang, T. Shi. New robust watermarking algorithm for vector data. *Wuhan University Journal of Natural Sciences*, 15(5): 403–407, 2010.
- [33] Y. Chauhan, P. Gupta, K. Majumder. Digital watermarking of satellite images. In proceedings of the 3rd Indian Conference on Computer Vision, Graphics and Image Processing, 2002.
- [34] A. Ho. Robust copyright protection of satellite images using a novel digital image-in-image watermarking algorithm. In proceedings of the International Geoscience and Remote Sensing Symposium, 2001.
- [35] B. Kumari and V. Rallabandi. Modified patchwork-based watermarking scheme for satellite imagery. *Signal Processing*, 88(4): 891–904, 2008.
- [36] P. Zhu, F. Jia, J. Zhang. A copyright protection watermarking algorithm for remote sensing image based on binary image watermark. *Optik-International Journal for Light and Electron Optics*, 124(20): 4177–4181, 2013.
- [37] J. Ruiz and D. Megias. A novel semi-fragile forensic watermarking scheme for remote sensing images. *International Journal of Remote Sensing*, 32(19): 5583–5606, 2011.
- [38] M. Barni, E. Magli, R. Troia. Minimum-Impact-on-Classifer (MIC) watermarking for protection of remote sensing imagery. In proceedings of IEEE International Geoscience and Remote Sensing Symposium, 2004.
- [39] M. Barni, F. Bartolini, V. Cappellini, E. Magli, G. Olmo. Near-lossless digital watermarking for copyright protection of remote sensing images. In



proceedings of International Geoscience and Remote Sensing Symposium, 2002.

- [40] M. Barni, F. Bartolini, A. Piva, F. Salucco. Cartographic image watermarking using text-based normalization. In proceedings of IEEE 4th Workshop on Multimedia Signal Processing, 2001.
- [41] L. Li and S. Guang. A watermarking algorithm for remote sensing image based on DFT and watermarking segmentation. Journal of Advanced Material Research, 433-440: 2504–2508, 2012.
- [42] T. Hemalatha, V. Jeevivek, K. Sukumar, K. Soman. Robust watermarking of remote sensing images without the loss of spatial information. In proceedings of 10th ESRI India User Conference, 2009.
- [43] L. Jing, Y. Zhang, G. Chen. Zero-watermarking for copyright protection of remote sensing image. In proceedings of the 9th International Conference on Signal Processing, 2008.
- [44] S. Ziegeler, H. Tamhankar, J. Fowler, L. Bruce. Wavelet-based watermarking of remotely sensed imagery tailored to classification performance. In proceedings of IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 2003.
- [45] P. Hsu and C. Chen. Feature-based digital watermarking for remote sensing images. ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, 2012; 1: 473–478.
- [46] A. Ho, J. Shen, S. Tan. Character-embedded watermarking algorithm using the fast Hadamard transform for satellite images. In proceedings of SPIE Conference on Mathematics of Data/Image Coding, Compression, and Encryption V, with Applications, 2003.
- [47] X. Zhu and A. Ho. A slant transform watermarking for copyright protection of satellite images. In proceedings of the 4th Pacific Rim Conference on Multimedia, 2003.
- [48] A. Liu and G. Lv. Copyright protection of DEM by means of DCT-based digital watermarking. Geo-information Science, 10(2): 214–223, 2008.
- [49] M. Barni, F. Bartolini, E. Magli, O. Gabriello, R. Zanini. Copyright protection of remote sensing imagery by means of digital watermarking. In proceedings of SPIE Conference on Sensors, Systems, and Next-Generation Satellites, 2001.

Biographies



Sangita Zope Chaudhari received M.E degree in computer engineering from Mumbai University, India. Currently she is a Ph.D student in Centre of Studies in Resources Engineering at Indian Institute of Technology Bombay, India.

Her research interests include digital image processing, advanced database and information systems, and information security techniques.



Parvatham Venkatachalam received the M.Sc. and Ph.D. degree in mathematics from Indian Institute of Technology Bombay, Mumbai, India in 1972 and 1978 respectively.

Currently, she is a professor in Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, India.

Her research interests include development of GIS for natural and human resources applications, digital image processing of remote sensing satellite data, development of spatial decision support systems, data structure in spatial databases, spatial data mining and warehousing.

Dr. Venkatachalam is a founder member of Indian Society of Geomatics, Life Member of Indian Society of Remote Sensing and Life Member of Indian National Cartographic Association.



B. Krishna Mohan received the Ph.D. degree in electrical engineering from the Indian Institute of Technology (IIT) Bombay, Mumbai, India, in 1991. He is currently Professor and Head with the Centre of Studies in Resources Engineering, IIT Bombay.

His areas of research include satellite image processing and analysis, machine learning algorithms, and educational content development for remote sensing and image processing.

Dr. Mohan is a Life Member of the Indian Society for Remote Sensing and Indian Society for Geomatics. He was the recipient of the 2003 IETE M.N. Saha Memorial Award for Best Application Oriented Paper in IETE Journal of Research.







Multistage VQ based Feature Vector for Effective CBIR

V. R. Khapli^A, A. S. Bhalchandra^B

A. Principal, K K W Women's Polytechnic, Amrutdham, Panchavati, Nashik and Research scholar at GEC Aurangabad,

B. Professor E and T/C, Government Engineering College, Usmanpura, Aurangabad, M.S. India
[vrkhopli@yahoo.com, asbhalchandra@gmail.com]

Abstract

Vector quantization (VQ) technique provides an effective codebook as representation of an image. This codebook is also very much useful as image feature from the view point of content based image retrieval (CBIR). Thus, designing an efficient codebook is the most important task in VQ. There are already several algorithms published on how to generate a codebook. However these codebooks when used as feature vector for CBIR still lacks the effectiveness as far as retrieval accuracy is concerned. To improve the retrieval accuracy, more effective feature vector is needed. To extract effective features for improving CBIR performance, a two pass VQ (MVQ) based FV technique is proposed in this paper and tested on Wang's generic color image datasets. Also a novel framework for Multistage VQ (MSVQ) based CBIR using distortion value in every code vector is proposed for extracting effective FV. The proposed systems proved to be promising for improving retrieval accuracy for generic database search.

Keywords: *Vector Quantization, Content Based Image Retrieval, Multistage Vector Quantization*

Nomenclature:

CBIR	Content Based Image Retrieval
C_i	Image codebook
C_Q	Query image codebook
$d(X, Y)$	Distance between vector X and vector Y
E_i	Error vector at i^{th} stage
FV	Feature Vector
Q1	Quantizer stage 1
Q2	Quantizer stage 2
SM	Similarity Metric
MVQ	Multiple VQ
MSVQ	Multistage VQ
SVQ	Single stage VQ
VQ	Vector Quantization
X	Set of training vectors
X'	Quantized vector

1. Introduction

Content-based image retrieval (CBIR) is a technique that helps to organize digital pictures by corresponding visual content. Many CBIR systems have been developed, but the problem of retrieving images on the basis of corresponding pixel content still remains a challenge [13, 15 and 21]. Vector quantization is a popular and powerful data compression technique for image coding [5]. It is also widely used for feature vector extraction in CBIR system [1, 4, 12, and 15]. The algorithms are proposed for fast VQ codebook generation [7]. The performance of a vector quantizer has reached the theoretical upper limit. However as the vector dimension increases the complexity involved also increases. Hence it has become very difficult to use high dimensionality VQ in practice. Generally, the encoding complexity increases exponentially with rate or dimension. Several VQ techniques have been proposed to solve the complexity issue. Some of the examples are: transformed VQ, shape gain VQ, tree structured VQ, multistage VQ, hierarchical VQ, predictive VQ, interpolative VQ. The codebook generated by VQ is suitable as feature vector for the image and is used to represent the image.

It is noticed that retrieval time is reduced due to use of reduced sized image for VQ based feature extraction without compromising retrieval accuracy [16]. Earlier two approaches are proposed and tested for making CBIR more effective. They were, VQ codebook initialization with reduced size image [17] and Fast encoding using Transformed domain VQ with an approximate image derived from low frequency sub image [18]. Both the techniques improved the retrieval time but there remains a scope for improving retrieval accuracy.

Multistage VQ has been used for image compression [2, 14]. MSVQ is also proposed for low bit rate video compression along with Multiwavelets [3]. This has given the motivation to try MSVQ based FV extraction for CBIR. To improve the accuracy of VQ based CBIR, two schemes are proposed in this paper, one using Multiple VQ (MVQ) and second using Multistage VQ (MSVQ), for image coding. This technique enables VQ of image blocks with tolerable encoding complexity and used as a feature vector for CBIR. MSVQThe performance of the proposed



schemes is compared with that of CBIR system using single stage VQ.

The remainder of the paper is organized as follows: Section(2) focuses on CBIR using VQ, Section(3) describes the proposed techniques, section (4) and (5) display sample outputs of proposed schemes, section (6) compares performance of proposed schemes with that of VQ based CBIR. Conclusions are stated in section (7) and references are listed in section (8).

2. CBIR using VQ

Vector Quantization (VQ) is one of the simple and effective techniques of feature vector extraction for CBIR. Refer figure 1. VQ requires decomposition of image into non overlapping sub image blocks. These blocks are termed as vectors. From all these vectors a set of representative image vectors is selected to represent the entire set of image blocks. The set of representative image vectors is called a codebook and each representative image vector is called code word. This codebook is the unique representation of an image and can be used as image signature or Feature vector of the image for CBIR [1].

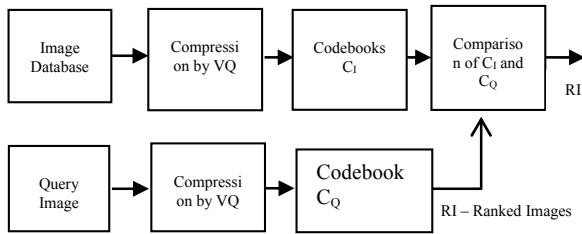


Figure 1. CBIR using VQ

Even though information is lost due to the compression, image retrieval based on VQ data not only provides information on the color content, but also on the spatial information (encompassing textural and shape attributes) of the image, which is due to the image being divided into blocks and the blocks coded as a whole. The most important task in VQ is of designing an efficient codebook. There are already several algorithms [4] published on how to generate a codebook. The *LBG algorithm* [20] is known as the Generalized Lloyds Algorithm (GLA). It is the most cited and widely used algorithms for designing the VQ codebook. It is the starting point for most of the work on vector quantization.

VQ based image classification and retrieval has been proposed in past. In VQ based method, quantization is performed on image blocks instead of single pixels. VQ based techniques are effective for Content Based Image Retrieval of compressed as well as uncompressed images [16]. Transformed VQ based technique is proposed for effective CBIR [18].

In this paper, two new VQ based techniques for Feature Vector generation are proposed. One using multiple VQ (MVQ) and the second is using Two stage VQ (MSVQ).

2.1 Multistage VQ

Multistage VQ is also known as cascaded VQ. It is also referred as residual VQ [5]. Here the basic idea is to divide the encoding task into successive stages. The first stage performs relatively crude quantization of the input vector. The second stage quantizer operates on an error vector between the original i.e. initial (randomly chosen from the image vectors) and quantized first stage output. The quantized error vector provides a second approximation to the input vector and hence leads to a refined or more accurate representation of the input. A further refinement may be achieved by a third stage quantizer and so on. Schematic of multistage VQ is shown in figure 2 where X is a set of training vectors, CB_i is i^{th} codebook and e_i is error at i^{th} stage.

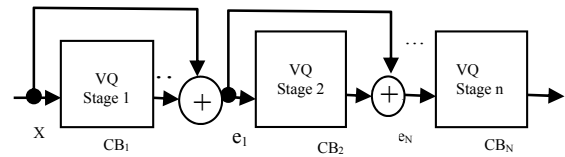


Figure 2. Multistage VQ

Figure 3 shows a special case of MSVQ i.e. two stage VQ.

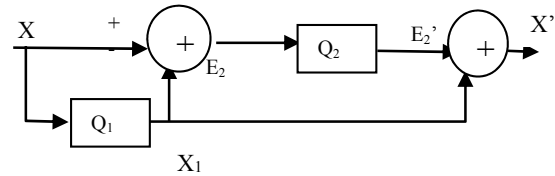


Figure 3. Two stage VQ

In two stage VQ the overall approximation X' to the input X is formed by summing the first and second stage approximations, X_1' and E_2 . From figure 2 it is seen that the input-output error is equal to the quantization error introduced by the second stage, i.e. $X - X' = E_2 - E_2'$

Therefore,

$SNR = SNR_1 + SNR_2$; where SNR_i is the signal to noise ratio in dB for the i^{th} quantizer.

Thus, a two stage quantizer has advantages of reduction in codebook size of each stage and lower complexity of searching.

The advantages are at the cost of reduction in overall SNR. However for CBIR it does not matter as we will be using the codebook as feature vector i.e. image signature and not reconstructing the image from the codebook.

In this paper, a two-stage vector quantizer is implemented. The input vector is quantized by the initial or first stage vector quantizer denoted by VQ_1 whose code book is $C_1 = \{c_{10}, c_{11}, \dots, c_{1(N_1-1)}\}$ with size N_1 .

The quantized approximation x_1' is then subtracted from x producing the error vector. This error vector is then applied to a second vector quantizer VQ_2 whose code book is $C_2 = \{c_{20}, c_{21}, \dots, c_{2(N_2-1)}\}$ with size N_2 yielding the quantized output. The encoder transmits a pair of indices specifying the selected codeword for each stage and the task of the decoder is to perform two table lookups to



generate and then sum the two code words. In fact, the overall codeword or index is the concatenation of code words or indices chosen from each of two codebooks. Thus, the equivalent product codebook can be generated from the Cartesian product $C1 \times C2$. Compared to the full-search VQ with the product codebook C , the two stage VQ can reduce the complexity from $N = N1 \times N2$ to $N1 + N2$.

3. Proposed Schemes for Feature Vector Extraction

To improve the performance of the Spatial Domain CBIR using VQ two novel schemes are proposed for FV extraction A. using multiple VQ (MVQ) and B. Using Multistage VQ (MSVQ).

3.1 CBIR using MVQ

A two stage VQ is used for Feature Vector (FV) extraction. Initially a codebook is obtained by using GLA (Generalized Lloyds Algorithm) as a first stage of VQ. Then this codebook is used as initial codebook and GLA is applied again to generate the final FV as the output of second stage VQ.

The proposed algorithm is represented in figure 4.

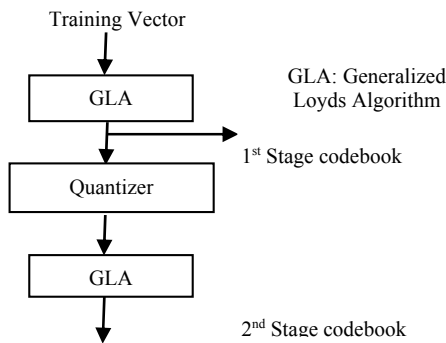


Figure 4. Codebook generation using Multiple VQ (MVQ)

All images are resized to 256 x 256 for uniformity in FV size. Image is transformed from RGB color space to the perceptually uniform *CIE LUV* color space and decomposed into non overlapping blocks. Any N vectors are picked randomly. Clusters are formed using GLA and K centroids are generated. These k centroids are used as k code vectors to form a codebook for the image. (VQ stage 1). This codebook is used as initial codebook and GLA is used again to get a new codebook with k vectors. (VQ stage 2). Resulting codebook is saved as image feature vector (FV) i. e. image signature. For all images in database, FVs are generated. This is preprocessing stage.

Similarly for query image as well, same steps are implemented and Query FV is generated.

For Retrieval, query FV is compared with all FVs in image database using Euclidian Distance as SM and the distance between them is calculated. Images are displayed in rank order of their distances.

3.2 CBIR using Multistage VQ (MSVQ) based Feature Vector

In the second approach for FV generation, a two stage VQ based FV is proposed for refining the initial codebook and the results are quite encouraging.

The proposed algorithm is as follows:

All images are resized to 256 x 256 and transformed to perpetually uniform *CIE LUV* color space. The three channels are then separated. For each channel then mean and variance are calculated by block processing as code vectors. Then applying k means clustering technique for vector quantization three code vectors are obtained for three channels ($m1$, $m2$ and $m3$ resp.)

Error vector is obtained by subtracting average distortion of all code vectors from every element of mean code vector 1 for each channel. Again VQ is performed on error vectors to get Feature vectors of the image from mean code vector 2 of each channel, that is nothing but the result of second stage VQ.

Similarly feature vector based on variance is obtained for three channels separately. Then the final feature vector is obtained by taking mean and variance vectors of stage 1 as well as stage 2 for all the three channels.

In this way all images in database are represented by their FVs as their image signatures.

A query image is also represented by its feature vector as image signature in the same manner.

Retrieval process is same using Euclidian Distance as Similarity Measure (SM) and then displaying the images in rank order of their SM.

The sample output of the proposed systems is displayed in figure 6.

3.3 Dataset used

Tests have been performed on the Wang's database of 1000 JPG images [19]. Ground truth on the image set is also defined. Images can be ranked in ascending order of the calculated distance. A subset of the images is assigned as query images and for each of these a series of correct matches that an ideal image indexing system would retrieve is specified. The datasets and the ground truth files are available at <http://wang.ist.psu.edu>.

3.4 Distance Measure

It is the similarity measure between two images. After indexing database images using extracted FVs, it is essential to determine similarity between Query FV and FVs of target images in the database. Similar images should have smaller distance between them. Images are ranked in order of their similarity measure. Many similarity measures like Euclidian Distance, city block distance, Canberra Metric, Chi-Square distance etc. are defined for CBIR [8, 10]. Euclidian distance is used as distance measure in the proposed schemes as it yields good accuracy and it is the most commonly used distance measure for CBIR. Euclidean distance is calculated as



defined below. The distance 'd' between the query image Q and an image V can be calculated as follows:

Let $X = \{x_1, x_2, \dots, x_k\}$ be a training vector and $d(X, Y)$ be the Euclidean distance between any two vectors X and Y. Then, $d(X, Y)$ is given by,

$$d(X, Y) = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (1)$$

Where x_i is i^{th} vector in query image codebook,
 y_i – i^{th} vector in target image codebook, and
 k – number of code vectors.

Images can be ranked in ascending order of the calculated distance. The smaller the calculated distance between two images, the more similar the two images.

3.5 Implementation

The proposed system is implemented using MATLAB 2007b on Intel Core(TM) i3 -2310M processor with 2GB RAM and 32 bit O.S. and 2.10 GHz clock.

4. Output of CBIR-VQ and CBIR-MVQ

The output of the system for sample query images with SVQ and MVQ based FV is shown in figure 5.

304 MVQ



304 SVQ



400 MVQ



400 SVQ



Figure 5(contd.). Sample Output of CBIR using MVQ and SVQ technique for two queries.

Query: Bus 305



Query Bus 304



Query Rose 600



Query image horse 700



Figure 5. Sample Output of CBIR using MVQ and SVQ technique for two queries

Figure 6. Sample Output of CBIR using a novel MSVQ based FV



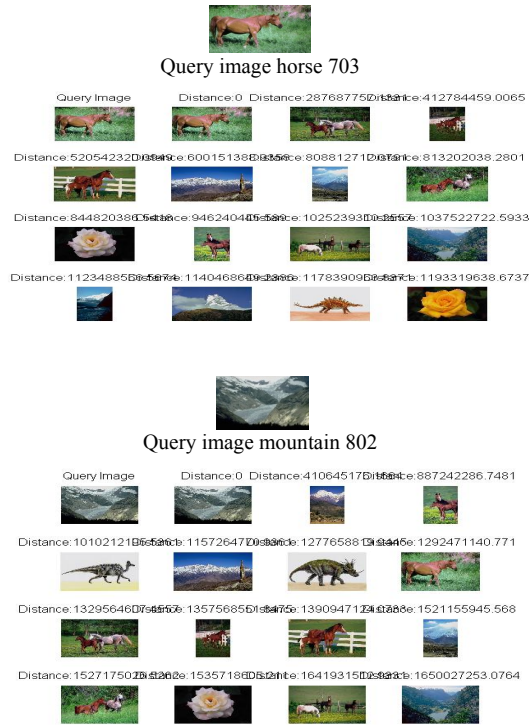


Figure 6(contd.). Sample Output of CBIR using a novel MSVQ based FV

5. Output of CBIR-MSVQ

The sample output of the CBIR system with MSVQ based FV is shown in figure 6 for various query images.

6. Performance Analysis

Performance of proposed system is evaluated in terms of retrieval accuracy

The most common performance measures of any CBIR system are Precision and Recall.

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (2)$$

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total relevant images in the database}} \quad (3)$$

Recall parameter is used to determine accuracy of CBIR for measuring its performance.

6.1 Performance of CBIR using MVQ

Table 1 shows comparison of the Retrieval Accuracy of the proposed scheme: CBIR using FV of Multiple VQ (MVQ) with that of CBIR using FV of Single VQ (SVQ). Figure 7 shows bar chart of improvement in the accuracy by deriving FV from Multiple VQ.

Table 1 Accuracy of CBIR with Multiple VQ and Single VQ (SVQ)

S.N	Query Image	Accuracy with SVQ in %	Accuracy with MVQ in %
1	304	57.14	85.71
Bus	305	42.85	100.00
	306	57.14	85.71
	310	71.42	100.00
Average % retrieval Accuracy		57.13	92.85
2	600	50.00	100.00
Flower	677	33.33	100.00
S.N	Query Image	Accuracy with SVQ	Accuracy with MVQ
Average % retrieval Accuracy		42.66	100.00
3	700	50.00	83.30
Horses	703	66.67	100.00
	702	50.00	100.00
	705	50.00	100.00
	701	50.00	83.33
Average % retrieval Accuracy		52.78	93.33
4	804	33.30	50.00
Mountain	805	33.30	66.60
	807	50.00	66.60
	803	40.00	60.00
Average % retrieval Accuracy		39.15	60.80
5	335	55.56	66.67
Mixed Q images not in the database	681	41.67	100
	410	70.00	100
	420	70.00	100
	635	50.00	100
	720	66.67	100
	615	58.33	100
	633	50.00	100
	417	70.00	100
	625	50.00	100
	681	41.67	100
	716	80.00	100
	710	80.00	100
	717	80.00	100
	345	85.71	100
Average % retrieval Accuracy		63.31	97.78



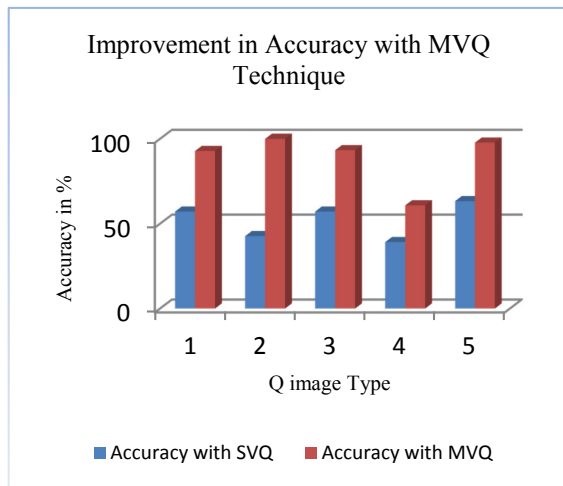


Figure 7. Comparison of Accuracies of the two CBIR systems

6.2 Performance of CBIR using MSVQ

Table 2 compares the average accuracy of CBIR-MSVQ with CBIR VQ. Figure 8 displays the comparison of performance of CBIR-VQ and CBIR-MSVQ in the form of bar chart.

Table 2 Average Accuracy with CBIR-VQ and CBIR-MSVQ

Query Image	Avg Accuracy of CBIR-VQ in %	Avg Accuracy of CBIR-MSVQ in %
Bus	57	81
Dianausor	70	65
Rose	43	97
Horse	53	63
Mountain	39	44

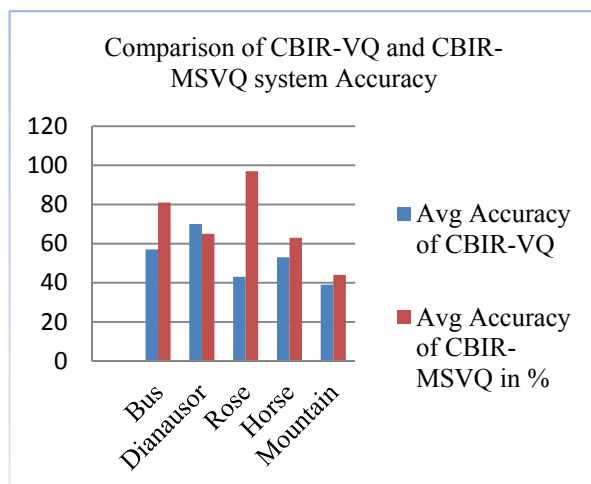


Figure 8. Comparison of Accuracies of the CBIR-VQ and CBIR-MSVQ systems

7. Conclusions and Discussions

FV extraction using VQ offers reduced complexity in CBIR. However the retrieval accuracy needs to be improved. Hence to make the VQ based FV more effective, a two stage VQ is proposed for FV extraction.

In CBIR-MVQ, FV is obtained by cascading two VQ stages. In the first stage the input is randomly selected vectors. The second stage operate on the output of quantized code vectors to generate better image signature as compared to signature generated by single stage VQ. It is observed by increase in the accuracy of CBIR-MVQ as compared with that of CBIR-VQ.

In CBIR-MSVQ quantized error vector that provides a second approximation of the input vector is used as image signature and hence leads to a refined or more accurate representation of the image. Thus MSVQ generates better image signature. As compared with single stage VQ based FV technique, Retrieval time increases (by 10 to 15 %) in MVQ and MSVQ based FV technique, but better image signature is achieved which improves the Retrieval Accuracy by 10-15%.

It is observed that adding more stages of VQ for FV extraction is not leading to improvement in its effectiveness. Hence two stage VQ is recommended for effective FV extraction.

Even though Euclidian distance is the most popularly used distance measure for CBIR, it is not the best in all cases as experimented by Kokre and al., as well as by M. Hatzigiorgaki and al., different Similarity Metrics may be tried to get even better accuracy using the MSVQ/MVQ based FV for CBIR.

8. References

- [1] A H Daptardar, J A Storer, "Reduced complexity Content Based Image Retrieval using Vector Quantization", Proceedings Data Compression Conference, IEEE Computer Society Press, 2006, pp 342-351
- [2] B. T. Hameed, "Design and Implementation of Multistage Vector Quantization Algorithm of Image compression assistant by Multiwavelet Transform", Diyala Journal for pure sciences, Vol: 6 No: 2, April 2010, pp19-31
- [3] E. Sankaralingam, V. Thangaraj, S. Vijayamani, N. Palaniswamy, Video Compression Using Multiwavelet and Multistage Vector Quantization , The International Arab Journal of Information Technology, Vol. 6, No. 4, October 2009 pp 385-411
- [4] G. Lu and S. Tang, "A Novel Image Retrieval Technique Based on Vector Quantization", Proceedings of International Conference on Computational Intelligence for Modeling, control and Automation, [1999], pp 36-41
- [5] Gersho and R. Gray, "Vector quantization and signal compression." Kluwer Academic Publisher, 1992
- [6] Hosam Khalil and Kenneth Rose, Predictive Multistage Vector Quantizer Design Using Asymptotic Closed-Loop Optimization, IEEE Transactions on Image Processing, Vol. 10, No. 11, pp 1765-1770
- [7] Jim Z. C. Lai, Y. C. Liaw, and Julie Liu, "A fast VQ codebook generation using codeword displacement,"



Pattern Recognition, vol. 41, no. 1, pp. 315-319, January 2008

- [8] Kokare M., Chatterji B. N., Biswas P. K., "Comparison of similarity metrics for texture image retrieval", TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region, pp 571- 575 Vol.2 ISBN: 0-7803-8162-9,
- [9] M. Satya Sai Ram, P. Siddaiah & M. Madhavi Latha, "Switched Multistage Vector Quantizer", Signal Processing: An International Journal (SPIJ) Volume (3) : Issue (6) 172-179
- [10] M. Hatzigiorgaki and A. N. Skodras , "Compressed Domain Image Retrieval: A Comparative Study of Similarity Metrics", SPIE Visual Communication and Image Processing, Vol. 5150, pp 439-448, 2003
- [11] P.V.N.Reddy, Dr. K.Satya Prasad, Multiwavelet Based Texture Features for Content Based Image Retrieval, IJCST Vol. 2, Issue 1, March 2011 pp 141-145
- [12] R. Nagendran and P. Arockia Jansi Rani, "A Novel FPCM based Vector quantizer codebook design for image compression in the wavelet packet domain", the international journal of multimedia and its applications, Vol2, NO. 1, Feb 2010, pp1-8
- [13] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age", ACM Computing Surveys, Vol. 40, No. 2, Article 5, Publication date: April 2008
- [14] S. Esakkirajan, T. Veerakumar, V. Senthil Murugan, P. Navaneethan , Image Compression Using Multiwavelet and Multi-stage Vector Quantization, Inter. Journal of Information and Communication Engineering 4:4 2008, pp 246-253
- [15] V R Khapli, A S Bhalchandra, "Compressed Domain Content Based Image Retrieval: State of the Art, Challenges and Open Issues", Proc. of World Academy of Science, Engineering and Technology, Bangkok, Thailand, Volume 36, December 2008 , pp 1335-1339, ISSN 2070-3740
- [16] V R Khapli, A S Bhalchandra, "Image Retrieval for Compressed and Uncompressed Images", IEEE Seventh International Conference on Information, Communications and Signal Processing (ICICS 2009), 8-10 Dec, Macau, China , pp 1-4
- [17] V. R. Khapli, A S Bhalchandra, "Spatial Domain Approach for Effective Image Retrieval using Thumbnails of Images", ICGST-GVIP Journal, Volume 11, Issue 3, June 2011 pp 9-17
- [18] V R Khapli, A S Bhalchandra, "Fast Encoding Technique of Transformed VQ for Effective CBIR", ICGST-GVIP Journal Vol. 12, Issue 2, August 2012 , pp 1-8
- [19] Wang's generic colour image database, <http://wang.ist.psu.edu/>
- [20] Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design, IEEE Transactions on Communications", 1980, vol. 28, pp. 84-95
- [21] Yong Rui and Thomas S. Huang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues". Journal of Visual Communication and Image Representation 10, 39-62 (1999)

Biographies



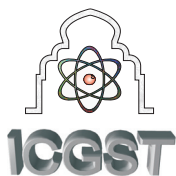
Vidya R Khapli is BE (Electronics and Power), (1985), from College of Engg, Amravati, M.S. India and M.Tech (Electronics) from VNIT, Nagpur, M.S. India, in 1996. Currently she is Principal at K K Wagh Women's Polytechnic, Nashik, MS, India. She is a research scholar pursuing for Ph.D. at Govt. Engineering College Aurangabad, M.S., India. Her areas of interest are Image Processing, Digital Signal Processing, Parallel Computing and Basic Electronics. Her research area is Content Based Image Retrieval.



Dr. Anjali Bhalchandra has received B.E. (Electronics & Telecommunication), M.E. (Electronics), Ph.D. (Electronics) from SGGS college of engineering, Nanded, Maharashtra, India. She is in the field of Engineering Education for last 29 years. Currently she is working as Head of E & T/C Dept. at GEC, Aurangabad (An Autonomous Institute of Govt. of India). Her research interests are in the area of image and signal processing with special focus on segmentation, CBIR techniques and Blind signal processing. She has more than 50 publications to her credit published at various international Journals, conferences and seminars. She has delivered several lectures on Image processing in Universities of India







Mutual Use of Derived Information from Radarsat-2 and ETM-8 Data for the Improvements of Feature Extraction

A. K. Helmy, Ayman H. Nasr and Safaa M. Hassan

National Authority for Remote Sensing and Space Sciences (NARSS), 23 Joseph Tito st., El-Nozha El-Gedida, P.O. Box: 1564 Alf-Mascan, Cairo, Egypt

akhelmy@narss.sci.eg, aymanasr@hotmail.com and safaamh2002@yahoo.com

Abstract

A new approach to improve the feature extraction and map updating of part of Halayib area located in south eastern desert of Egypt, using multisource data sets, is proposed in this paper. It investigates the potential of the joint use of derived information from both PolSAR and Optical images. As they refer to different properties of the observed scene, it's expected to complement each other. The approach is consisted of combining Radarsat-2 Pauli decomposition data, with ETM-8 principal component transformation. We used the bandlet transform to decompose the original images into a series of frequencies. It depends on the geometric flow of an image to construct the bandlet basis. The geometric flow takes advantage of the geometric regularity of the image structure and represents sharp image transitions such as edges efficiently in the output image. For producing the fused image, we proposed a fusion rule to first-rate the geometric flow of the source images, and then employed the m-PCNN to implement the fusion process with the bandlet coefficients. The new results were evaluated using qualitative and quantitative measures. They demonstrate improved interpretation, where PolSAR texture information enhances the fusion product. The proposed approach clearly states that, the combination of PolSAR and Optical derived data has potential to improve feature extraction and map updating in desert area.

Keywords: PolSAR, ETM-8, Data fusion, m-PCNN, Bandlet Transform.

Nomenclature

SAR	Syntactic Aperture Radar
PolSAR	Polarimetric SAR
ETM-8	Enhanced Thematic Mapper
m-PCNN	Multi-Channel Pulse Coupled Neural Network
PC	Principal Component.
[S]	Radar Scattering matrix.
MI	Mutual Information.
E	Entropy
Av Grad	Average Gradient.

1. Introduction

The processing of desert area is sometimes difficult to analyze using Polarimetric SAR (PolSAR) or optical data alone. This research investigates the applicability of PolSAR data products in how its combination with other optical derived data can introduce a new set of data for more feature extraction, improved interpretation and significant insight in desert areas. Radarsat-2 Pauli decomposition provides an excellent opportunity to research the potential of fully PolSAR. It was combined with the Enhanced Thematic Mapper(ETM-8) derived data; principal components transformation to create the new data set. While these two algorithms are rather well known, this example is given for the first time. The pixel value in optical imagery represents the reflectivity of the object in the spectral range of the detectors. In SAR imaging, a pixel value is the sum of the multiple coherent backscatters of the transmitted signal from the objects in the scene. In case of PolSAR imaging, the polarization information captured is largely uncorrelated with the spectral and intensity images. The intensity image provides information on materials in the scene while polarization measurements capture surface features, roughness, and shading, often uncorrelated with the intensity image [1].

The remainder of the paper is structured as follows: Section 2 presents the main concept of Polarimetric Syntactic Aperture Radar (PolSAR) imagery. Section 3 and 4 introduce a brief discussion of Pauli decomposition applied to PolSAR images and principal component analysis applied to ETM-8 multispectral imagery, respectively. Section 6 focuses on a detailed description of multi-channel Pulse Coupled Neural Network (m-PCNN); main structure and various parameters used to adapt the performance of the network. The study area and data acquisition of the satellite images are outlined in section 7. Section 8 depicts the full description of a proposed method for the data fusion accompanied with the interpretation of the main results. Finally, Section 9 summarizes the conclusions achieved by this study.



2. Polarimetric SAR

Polarimetric SAR (PolSAR) provides an enhanced capability for advanced target recognition in different remote sensing applications. It is implemented by alternatively transmitting horizontally and vertically polarized waves and receiving both polarizations. It records four channels: horizontal transmit-horizontal receive (HH), horizontal transmit-vertical receive (HV), vertical transmit-horizontal receive (VH) and vertical transmit-vertical receive (VV). Polarimetric sensors observe a scattering matrix that gives information about amplitude and phase of the backscattered signal. To identify the occurrence of the scattering mechanisms, the scattering matrix is decomposed [2]. We are going to use the Coherent (Pauli) decomposition since it is based on linear combinations of the PolSAR channels (HH, HV, VH and VV) [3] and preserves the polarization information (i.e., all information is maintained). In case of (incoherent) decomposition, due to local averaging, the phase information will be lost.

3. Pauli Decomposition

The Pauli decomposition is the most common SAR decomposition [4]. It is often employed to represent all the polarimetric information in a PolSAR image. It is a method for breaking the data down into components explaining surface scattering properties. This decomposition represents the scattering matrix [5] as three components representing single-bounce, double-bounce, and volumetric scattering mechanisms. In comparison to other coherent decomposition methods, the Pauli decomposition is excellent for exposing natural targets, but not ideal for highlighting man-made targets [4]. The scattering matrix [S] can be written as:

$$[S] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} = \alpha \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \beta \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \gamma \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (1)$$

$$\text{where } \alpha = \frac{(S_{hh} + S_{vv})}{\sqrt{2}}, \quad \beta = \frac{(S_{hh} - S_{vv})}{\sqrt{2}}, \quad \gamma = \sqrt{2}S_{hv}$$

The Pauli decomposition's dimensionality of three makes it simple to represent visually using the RGB color scheme [6].

4. Principal Component Transformation

The Principal Component Analysis (PCA) and the Intensity-Hue-Saturation (IHS) techniques are often used to fuse optical and radar images [7]. In our study, we have used the PCA transform. It aims at creating new uncorrelated channels from initial observations. The covariance matrix and its diagonalization were computed by finding corresponding eigen values and eigenvectors. The fourth, fifth, and the eighth components of the PCA transform were used in this study. Finally, the enhanced output was retrieved by an inverse transformation. The PolSAR Pauli decomposition and the optical PCA images were fused together. The fusion was conducted based on bandlet transform and the multi-channel model of Pulse Coupled Neural Network (m-PCNN).

5. The Bandlet Transform

When dealing with PolSAR images, the textural information plays an important role in information fusion. The textural structures within these images are mostly irregular or diffuse edges. In the fusion process, it is mandatory to have a strong knowledge about the textural structure of the images, define changing zones and give important tracks to exploit data fusion rules. Image transformation usually tackles the problem of image texture from the image geometry point of view. There are several image transformation techniques in literatures such as wavelet, contourlet, and bandlet transforms. The bandlets are orthonormal basis [8]; all vectors in the set are mutually orthogonal and of unit length. They use adapted wavelet basis to perform a transform on functions defined as smooth functions on smoothly bounded domains. The bandlets have great ability to maintain the textural information content of the surfaces. They use an adaptive segmentation process to segment an image into several blocks with different sizes such that each block is regular in some manner.

In the segmentation process, the image is firstly divided into four blocks with equal sizes. Each block is examined from the variance point of view. If the block shows low variance value, this block is assigned as regular block. Else the block is further divided into four sub-blocks. The quad tree structure is used in the segmentation process. Another factor introduced in the segmentation process, namely geometric flow, which can identify the direction of variation in the neighborhood of a candidate pixel; most commonly the variation could be in horizontal or vertical directions. According to this process, each sub-block could be assigned to one of three groups: Regular macro-block (R), the whole block is uniform, Horizontal edge macro-block (H), there is only one single horizontal edge in the sub-block, and Vertical edge macro-block (V), there is only one single vertical edge in the sub-block [9]. Finally, bandlet basis is calculated according to the regularity and geometric flow of each sub-block.

The operational procedures to construct the bandlet basis are as follows:

Let: S represents the image domain and $\tau(x_1, x_2)$ is a vector field represents the geometric flow that defines the direction of variation in the neighborhood (x_1, x_2) . Where the geometric flow is the slope of the vector field:

If the flow is assigned to parallel horizontally i.e. $x_2 = 0$ then $\tau(x_1, x_2) = \tau(x_1)$.

If the flow is assigned to parallel vertically i.e. $x_1 = 0$ then $\tau(x_1, x_2) = \tau(x_2)$.

Calculate the flow line as:

$$g(x_1) = \int_0^{x_1} g'(x) dx \quad (2)$$

Where $g'(x)$ is the slope of the geometric flow vector $\tau(x_1, x_2)$.

6. The Multi-channel Pulse Coupled Neural Network

PCNN is a simulation of biological neural network; basically it models a visual cortex of the brain which is responsible for processing visual information [10]. It is a two-dimensional neural network extended to be multi-



channel neural network (m-PCNN). Recently PCNN shows potential progress in the field of image processing such as image fusion, image segmentation, image enhancement and pattern recognition [11]. For variations in input patterns, PCNN shows independence of geometry, can bridge minor intensity and has robustness against noise. Usually PCNN is used in image processing community through specific structure as follows:

PCNN consists of many neurons, and the number of neurons is equal to the number of pixels in the input image. Each pixel of an input image is represented by solely neuron. It consists of three parts: input part, linking modulation, and pulse generator. The input part receives two types of information: Candidate pixel information (color, texture ...) transferred to the corresponding neuron by external stimulus (through linking channel) and Pixels information transferred between neighbored neurons by local stimulus (through feeding channel). It is clear that the feeding channels have a slower response time than the linking channels. The feeding and linking outputs are given by:

$$\begin{aligned} F_{i,j}[n] &= \exp(-\alpha_F) F_{i,j}[n-1] + V_F \sum_{k,l} m_{ijkl} Y_{i,j}[n-1] + S_{i,j} \\ L_{i,j}[n] &= \exp(-\alpha_L) L_{i,j}[n-1] + V_L \sum_{k,l} w_{ijkl} Y_{i,j}[n-1] + S_{i,j} \end{aligned} \quad (3)$$

Where F and L are the feeding and linking process, m and w are constant factors represent the strength of synaptic weights. S is the input stimulus. V_F , V_L are normalizing constant; α_F , α_L are time constants such that $\alpha_L > \alpha_F$.

The role of linking and modulation is to mix the outputs of linking and feeding channels through iterative computation and accumulates the resultant until it reaches a pre-specified threshold and it is given by:

$$U_{i,j}[n] = F_{i,j}[n](1 + \beta L_{i,j}[n]) \quad (4)$$

Where $U_{i,j}$ is the internal state of the neuron and β is bias constant.

The role of pulse generator is to generate pulses based on its input state and the threshold value. These output pulses contain information of the different input images and the threshold value is calculated as:

$$T_{i,j}[n] = \exp(-\alpha_T) T_{i,j}[n-1] + V_T Y_{i,j}[n] \quad (5)$$

Where V_T and α_T are normalized and time constant.

The output of the pulse generator is defined in the following equation:

$$Y_{i,j}[n] = \begin{cases} 1, & U_{i,j}[n] > T_{i,j}[n] \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

Many parameters are included in PCNN which have direct effect on the performance of the network. Choice of suitable values of these parameters depends mainly on the application. Here we will introduce a brief explanation about the function of these parameters and their selected values:

	Parameters	Descriptions
Feed channel	$\alpha_F = 0.1$	Controls the rate of decay of the feeding channel. Larger value leads to faster decay.
	$V_F = 0.5$	Represents the amount of interference between neighbored pixel and candidate neuron. Higher value means more influence from the surrounding neurons.
	$W = \begin{bmatrix} 0.5 & 1 & 0.5 \\ 1 & 0 & 1 \\ 0.5 & 1 & 0.5 \end{bmatrix}$	Refers to the size of the area which consists of the neighboring pixels of the corresponding pixel in the input image. It determines the synaptic weight strength.
Link channel	$\alpha_L = 1.0$	Controls the rate of decay of the linking channel. Larger value leads to faster decay.
	$V_L = 0.2$	Represents the amount of interference between neighbored pixel and candidate neuron. Higher value means more influence from the surrounding neurons.
	$M = \begin{bmatrix} 0.5 & 1 & 0.5 \\ 1 & 0 & 1 \\ 0.5 & 1 & 0.5 \end{bmatrix}$	Refers to the size of the area which consists of the neighboring pixels of the corresponding pixel in the input image. It determines the synaptic weight strength
Pulse Generator	$\beta = 0.1$	Considered as weighting factor of the linking channel in the linking and modulation activity. Higher value of β indicates more influence from the linking channel. The value of β could be the same for all neurons or each neuron may have its own value.
	$\alpha_T = 1.0$	Determines at what time the neuron should fire. It represents the rate of decay of the threshold in different iteration. Smaller value of α_T makes the network takes much time to fire.

PCNN has been used for integrating multiple images from different sources, Wang and Ma used a single layer of PCNN to fuse different medical images [12, 13]. Miao designed an adaptive linking coefficient for PCNN used in multi-resolution image fusion [14]. Others introduced the principal of combining PCNN with multi-layer decomposition for data fusion [15, 16]. Recently, multiple layers of PCNNs working in parallel are introduced in the field of image processing [11]. The main problem of PCNN is tuning the network parameters, since there are a lot of network parameters needed to be adjusted based on the nature of the problem, moreover the



limited capability of PCNN to handle a few numbers of inputs [17]. Undoubtedly, for multiple-image fusion, it is an obstacle to have one stimulus for each neuron in case of PCNN. Therefore, the new improved model m-PCNN is used [12], as shown in figure 1.

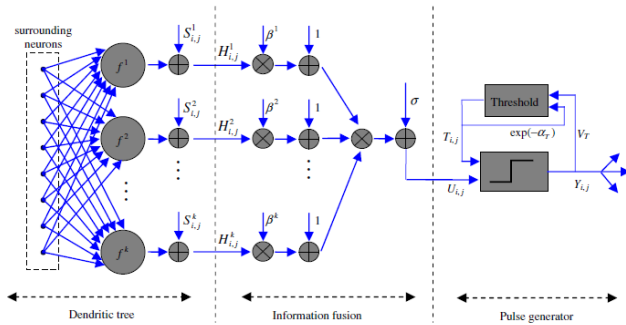


Figure 1. The improved model of m-PCNN (©Wang and Ma, 2008)

m-PCNN shared with the original PCNN in the neuron structure. Each neuron contains three main parts: “dendritic tree, information fusion, and pulse generator”. Dendritic tree receives two sort of information; one of the external stimulus and the other from the adjacent neurons. All input information is merged through information fusion place. The pulse generator generates the output pulse. Compared with the original model, m-PCNN introduce a dynamic way to adjust the number of inputs to suit the desired problem. All inputs can be input to the model at the same time.

7. The Study Area and Data Acquisition

The study area is part of Halayib area located in south eastern desert of Egypt, geographically situated at the coordinate of $22^{\circ} 29'$ to $22^{\circ} 09' N$, and $35^{\circ} 43'$ to $36^{\circ} 03' E$, as shown in figure 2.

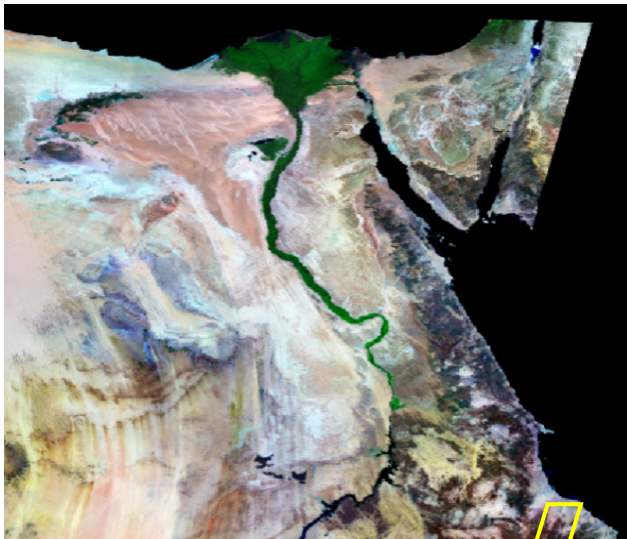
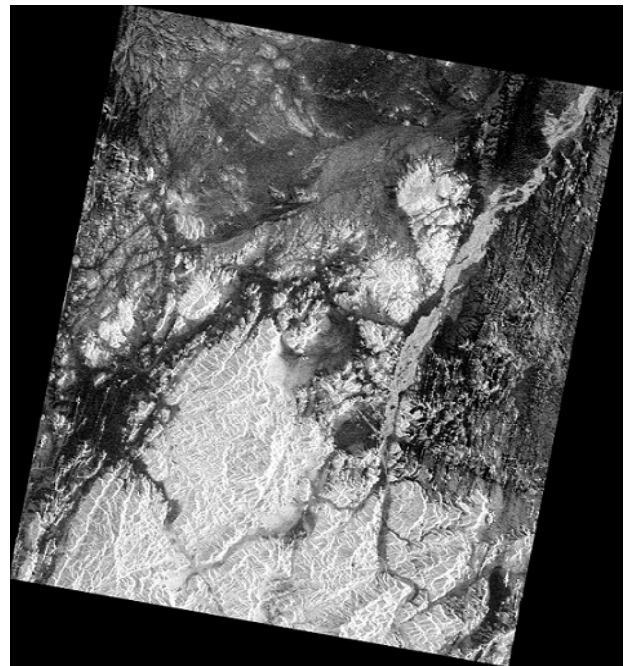


Figure 2. The study area (yellow rectangle)

Radarsat-2 PolSAR image, 8 m resolution acquired on 2013 and the multi-spectral bands of Landsat ETM-8 image, 30 m resolution acquired on 2014 were used in conducting this study, as shown in figure 3. The Software's used were the Polarimetric SAR Data Processing and Educational Toolbox; (PolSARpro) and the Matlab 2014a.



(a)- ETM-8 Image



b)- Radarsat-2 Span Image

Figure 3. (a, b) Data Used of the Study Area

8. The Proposed Method

In our proposed method, bandlet transform is used as a multi-scale decomposition tool for the input images, while m-PCNN is implemented for the image fusion. Bandlet transform has a great capability to extract the features of the original images, such as edges and texture which enhances a fusion process. The operational procedure can be outlined as follows:

- i) The ETM-8 and the PolSAR images are perfectly co-registered using different GCP points with sub-pixel precision. Then they were resampled to 15m using the nearest neighbor technique. [18]



summarizes the co-registration techniques of the optical and radar images.

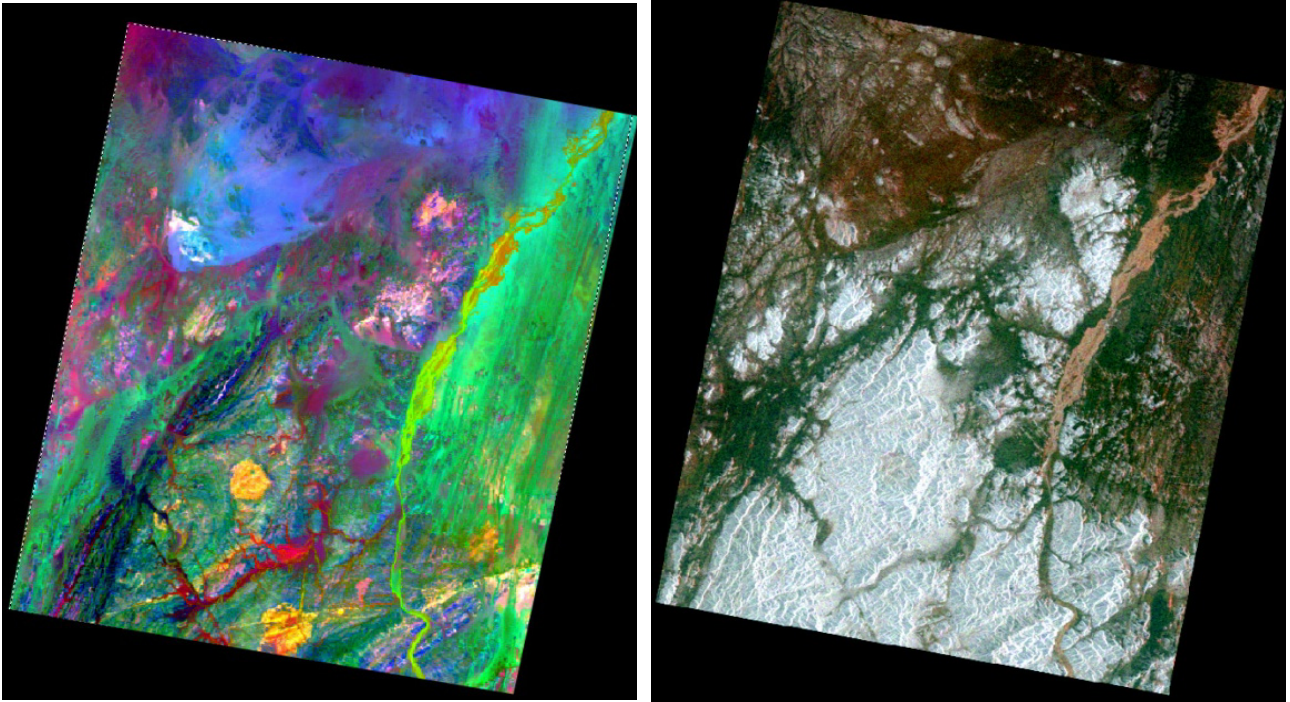
- ii) The ETM-8 image is transformed into the PCA components as shown in figure 4.
 - iii) The Pauli decomposition of the PolSAR image is performed as shown also in figure 4.
 - iv) The bandlet transform is applied to these two input images. They are segmented into several blocks with different sizes, such that each segmented block is either regular, having horizontal regular edges or vertical regular edges.
1. We calculated the Geometric flow (G) which is defined as the direction of variation in the neighborhood. Three types of variations are introduced here, regular variation (there is no variation in this region), horizontal variation (the variation is along horizontal direction) and the vertical variation (the variation is along the vertical direction).
 2. We also calculated the bandlet coefficient (C).

- v) The fusion rule is applied to the geometric flow (G) of both images. The fused geometric flow (G_F) at location (i, j) can be calculated by:

$$G_F(i, j) = \frac{W_1 G_1(i, j) + W_2 G_2(i, j)}{2} \quad (7)$$

The values of w_1 and w_2 are calculated according to the type of variations. If both regions are uniformly regular, then $w_1 = w_2 = 1$ (simple average weight). Else the weights are equal to the slope of variations.

- vi) The bandlet coefficients (C) of the input images are fed to the m-PCNN for the coefficients fusion process as shown in figure 5.
- vii) The Bandlet inverse transform is applied and the final output image was rescaled to the dynamic range (0-255).



(a) PCA of the ETM-8 (b) Pauli decomposition of the Radarsat-2

Figure 4. The Data Products Used in this Study

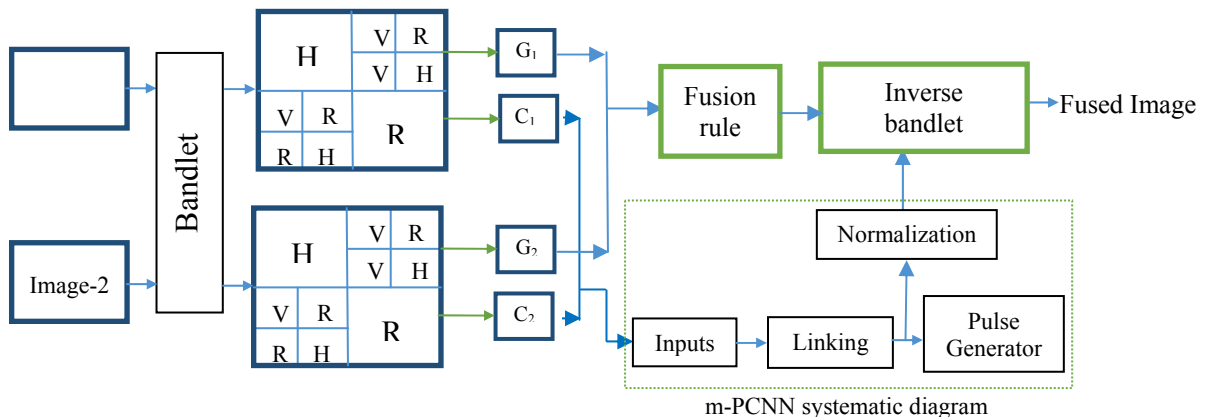


Figure 5. Proposed fusion framework



After applying the previous procedures we got the final output fused image as depicted in figure 6.

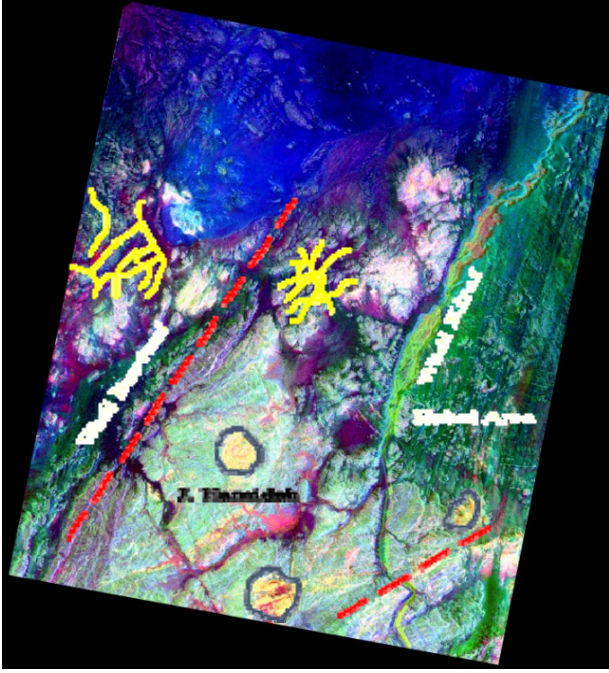


Figure 6. The output fused image

The performance of the proposed method was both qualitatively and quantitatively assessed. For the qualitative assessment; from the previous figure and by visual inspection, the following geological interpretation and discrimination could be deduced:

- 1- The subsurface drainages beneath the Quaternary sand plain (violet and dark colors) at the middle and the western corner of the image (yellow lines).
- 2- The Quaternary Wadi deposits, N-S trending, filling Wadi Kiraf (green color).
- 3- The rounded granitic plutons intrusions at the southern part of the image (light yellow - Blue circles).
- 4- The NNE-SSW parallel sand dunes similar dykes' east Wadi Kiraf- Hubal area (dark green).
- 5- The NE-SW trending faults at the SE corner and the western part of the image (red dashed lines).

The aridity of the sand and soils of the study area permitted radar subsurface penetration and the appearance of features beneath the sand dunes. Subsurface structures, paleo drainage pattern and the lithology have been enhanced in the fused image. The quantitative assessment was performed by adopting the following frequently used measures [19]:

- 1-Mutual information (MI) basically computes the amount of information transferred from the source images to the fused image. MI is calculated as follows:

$$MI = \frac{\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{A,F}(i,k) \log \frac{P_{A,F}(i,k)}{P_A(i)P_F(k)} + \sum_{j=0}^{L-1} \sum_{i=0}^{L-1} P_{B,F}(i,k) \log \frac{P_{B,F}(j,k)}{P_B(j)P_F(k)}}{IE_A + IE_B} \quad (8)$$

where $P_{B,F}(j,k)$ and $P_{A,F}(i,k)$ are the normalized gray histogram between the source image A and the fused image F , the normalized gray histogram between the source image B and F , respectively.

- 2-The Metric $Q^{AB/F}$ [20] is similar to MI ; it measures the amount of edge information transferred from the source images to the fused image using Sobel edge detector:

$$Q^{AB/F} = \frac{\sum_{n=1}^N \sum_{m=1}^M (Q^{AF}(n,m)W^A(n,m) + Q^{BF}(n,m)W^B(n,m))}{\sum_{n=1}^N \sum_{m=1}^M (W^A(n,m) + W^B(n,m))} \quad (9)$$

Where the value of $Q^{AF}(n,m)$ represents the orientation preservation values and the edge strength; N and M are the image size. The values of $Q^{AB/F}$ range from 0 to 1, its higher value represents better edge preservation

- 3-The entropy is another evaluator for the proposed algorithm performance. It is defined as:

$$E = - \sum_{l=0}^L P(l) \log_2 P(l) \quad (10)$$

where $P(l)$ means the probability of the gray value l appearing in the image; L is the largest gray level and $L = 255$ if the number of the image gray levels is 256. The entropy describes the average information of the image. The larger the value is, the more information will be contained in the image.

- 4-The average gradient (Av-Grad) reflects the difference between image details and texture variations. Generally,

the greater the average gradient is, the sharper the image will be. Table 1 shows the output values after applying the previous measures.

Image \ Measure	Fused	Pauli	PCA
MI	0.84	-----	-----
$Q^{AB/F}$	0.885	-----	-----
Entropy	6.9	5.4	4.8
Av-Grad	7.53	6.02	5.46

Table 1: The output values of the previous measures

The transferred information from the input images; Pauli and PCA to the fused image are 84% and 88% as calculated by the MI and $Q^{AB/F}$ measures, respectively. The average information of the image represented by the entropy shows highest value (6.9) in the fused image compared to the other input images. Finally, the Av-Grad of the fused image demonstrates higher edge strength value (7.53) than Pauli and PCA images which depicts the substantial improvements of the proposed approach.



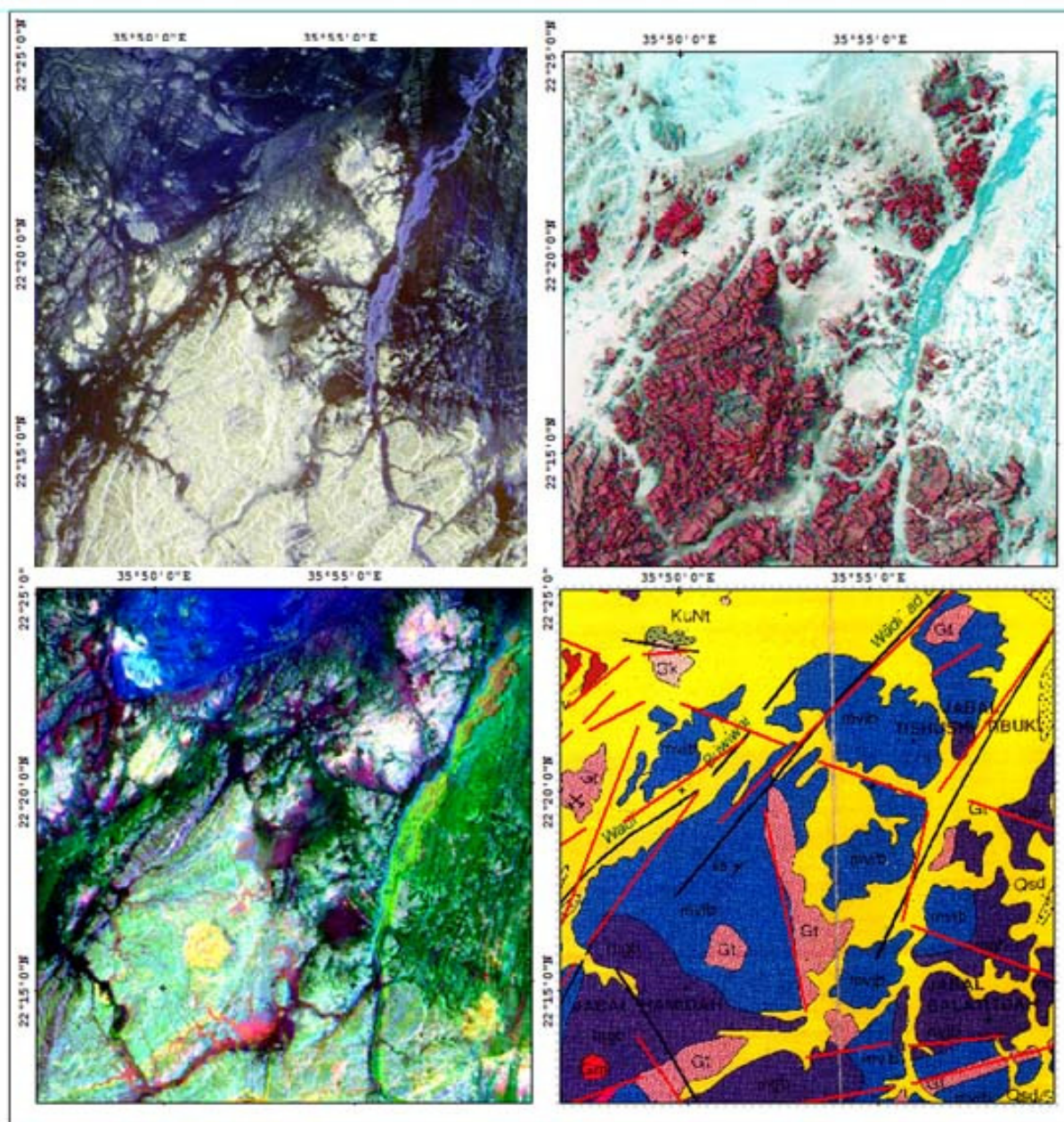


Figure 7. Zooming in part of the study area: (a) Pauli Decomposition of the PoLSAR image; (b) ETM-8 raw image; (c) The output fused image; (d) The geological map of the study areaImage-1

Figure 7 shows zooming in part of the study area. In addition to the previous interpretation, the output fused image (c) also discriminates the NE-SW trending sub-parallel alteration zones of talc chlorite schist (pinkish purple color) within the metavolcanic belt at Jabal Hamidah in the western part of the study area due to increasing of the back scattering as a source of information. These alteration zones are not clearly distinguished in both the radar and the ETM-8 raw data, as shown in figure 3. Therefore, geological map updating and new feature extraction could be achieved in a most revealing manner.

9. Conclusions

The motivation behind this research is to perform synergy between two complementary modalities; PolSAR and Optical derived data which is a new area in remote sensing. Finding new methods to fuse the

products of these data increased the information gain and improved the feature extraction that was not previously achieved with the help of any one of these modes of imaging individually. They showed the potential and benefits of combining Radarsat-2 Pauli decomposition and ETM-8 principal component transformation using m-PCNN and bandlet bases to better characterize and understand the study area. The new results were evaluated using qualitative and quantitative measures. The quantitative measures included; Mutual information (MI), the Metric ($Q^{AB/F}$), the entropy (E) and the average gradient (Av-Grad). These measures proved that the fused image was strongly correlated with the source images and provided substantial improvements. Therefore, the new approach can be more broadly applicable to provide an enhanced feature extraction, significant improvements in objects recognition and the creation of realistic geological maps in similar regions.



10. Acknowledgements

The authors would like to express their deepest thanks and appreciation to the Canadian Space Agency (CSA) for providing the Radarsat-2 image for this study under the agreement of the project ID5128. Also, Special thanks are extended to Prof. Dr. Mohamed Fouad, Geological mapping department at NARSS, for his interpretation of the output results.

11. References

- [1] Maitra S., 2013: "Analysis of Polarimetric Synthetic Aperture Radar and Passive Visible Light Polarimetric Imaging Data Fusion for Remote Sensing Applications". Ph.D. dissertation, The Chester F. Carlson Center for Imaging Science, College of Science, Rochester Institute of Technology, 189 p.
- [2] J. J. van Zyl, "Application of Cloude's target decomposition theorem to polarimetric imaging radar data," *Radar Polarimetry*, San Diego, CA, SPIE, vol. 1748, pp. 184–212, July 1992
- [3] Maitra, S., Gartley, M.G. and Kerekes, J.P., 2012: "Relation between degree of polarization and Pauli color coded image to characterize scattering mechanisms", *Proc. SPIE 8364, Polarization Measurement, Analysis and Remote Sensing X*, 83640F.
- [4] Zhang, L., Zhang, J., Zou, B. and Zhang, Y., 2008: "Comparison of methods for target detection and applications using polarimetric SAR image", *Progress in Electromagnetic Research, PIERS Online*, vol.4, no. 1, pp. 140–14.
- [5] Huynen, J. R., 1965: "Measurement of the Target scattering matrix", *Proceedings of the IEEE*, 53(8), 936.
- [6] Hall, W.D., 2012: "Exploration of Data Fusion between Polarimetric Radar and Multispectral Image Data", B.S. Thesis, Civilian Department of the Navy, California State University at Stanislaus, USA, 67 p.
- [7] Jinghui, Y., Jixian, Z., Haitao, L., Yushan, S. and Pengxian, P., 2010: "Pixel level fusion methods for remote sensing images: a current review". *IAPRS 38*, pp. 680–686.
- [8] Mallat, S., 1989: "A theory for multi-resolution signal decomposition: The wavelet representation", *Pami*, vol. 11, no. 7, pp. 674-693.
- [9] Le Pennec, E. and Mallat, S., 2005: "Sparse Geometric Image Representations with Bandlets," *IEEE Trans. Image Processing*, vol. 14, no.4, pp. 423-438.
- [10] Eckhorn, R., Reitboeck, H.J., Arndt, M. and Dicke, P.W., 1989: "A neural network for feature linking via synchronous activity: results from cat visual cortex and from simulations", in: *Models of Brain Function*, Cambridge University Press, Cambridge, UK, pp. 255–272.
- [11] Wang, Z., Ma, Y., Cheng, F., and Yang, L., 2010: "Review of pulse-coupled neural networks". *Image and Vision Computing*, vol. 28, pp. 5–13.
- [12] Wang, Z. and Ma, Y., 2008: "Medical image fusion using m-PCNN". *Journal of Information Fusion*, vol. 9, pp. 176-185.
- [13] Wang, Z. and Ma, Y., 2007: "Dual-channel PCNN and its application in the field of image fusion". *Proc. of the 3rd International Conference on Natural Computation*, vol. 1, pp. 755–759.
- [14] Miao, Q. and Wang, B., 2005: "A novel algorithm of multi-focus image fusion using adaptive PCNN", "A novel adaptive multi-focus image fusion algorithm based on PCNN and sharpness" *Proc. SPIE 5778, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV*, (704): doi:10.1117/12.603092.
- [15] Xu, B. and Chen, Z., 2004: "A multisensor image fusion algorithm based on PCNN". *Proc. of the 5th World Congress on Intelligent Control and Automation*, vol. 4, pp. 3679–3682.
- [16] Li, W. and Zhu, X., 2005: "A new image fusion algorithm based on wavelet packet analysis and PCNN", *Proc. of the 4th International Conference on Machine Learning and Cybernetics*, vol. 9, pp. 5297–5301.
- [17] El-taweel, G. S. and Helmy, A. K., 2013: "Image fusion scheme based on modified dual pulse coupled neural network", *IET Image Processing* 7(5), pp. 407–414.
- [18] Tupin, F., 2011: "Radar Remote Sensing of Urban areas", *Fusion of optical and SAR images*, Chapter 6, 'editions Springer'.
- [19] Qu, G.H., Zhang, D.L., and Yan, P.F., 2002: "Information measure for performance of image fusion", *Electron. Lett.* vol. 38, pp. 313–315.
- [20] Petrovic, V. and Xydeas, C., 2000: "On the effects of sensor noise in pixel-level image fusion performance". In: *Proc. of the Third International Conference on Image Fusion*, vol. 2, IEEE Press, pp. 14–19.



Biographies



Dr. Ayman Nasr is Associate Research Professor in National Authority for Remote Sensing and Space Sciences. He is currently the Head of the Data Reception, Analysis and Receiving Station Affairs Division. He received his B.Sc., M.Sc. and Ph.D. in Electronics and Communications from

Faculty of Engineering, Cairo University. His research expertise includes image processing and GIS in different fields of Remote Sensing. He has published 44 scientific papers in various national and international academic journals and conference proceedings. He has contributed in publishing 4 atlases and participated in more than 50 projects and researches related to remote sensing fields. He supervises and graduated several Masters and PhD students. Dr. Nasr is currently a member of the Editorial Board of the Egyptian Journal of Remote Sensing and Space Sciences. He is also a member of the Egyptian Society of Remote Sensing, GIS and Space sciences.



Dr. Ashraf Helmy is Associate Research Professor in the National Authority for Remote Sensing and Space Sciences. He is currently the Head of the Space Sciences and strategic studies Division. He has got his Ph.D. in information technology from Cairo University, 2005. His area of interest includes

signal and image processing.



Dr. Safaa Hassan has got her Ph.D. in Geology, GIS and Remote Sensing. She is currently the Head of Geographic information systems department of Data Reception, Analysis and Receiving Station Affairs Division, at the National Authority for Remote Sensing and Space Sciences. She has Broad experience in the analysis of space

and airborne sensor collected data, LIDAR data processing, digital image processing, and GIS, for remote sensing applications.







Classification of Crops using Polarimetric SAR Parameters based on Scattering Mechanism

Hind H. Zeyada¹, Mohammed M. Ezz², Ayman H. Nasr¹, Hany M. Harb²

¹ Data Reception, Analysis and Receiving Station Affairs (DRARSA),

National Authority for Remote Sensing and Space Science (NARSS) Cairo, Egypt.

² Faculty of Engineering, Al-Azhar University, Cairo, Egypt.

Abstract

The contribution of polarimetric data has a significant role in Land Cover (LC) classification. Polarimetric Synthetic Aperture Radar (SAR) data describes radar scattering mechanism quantitatively. In this paper, we use a few parameters to explore their potential in defining five crops and land cover surfaces: rice, maize, grape, cotton and urban, using Wishart classification. This classification is based on parameters resulting from decomposition of the coherency matrix for each pixel. The parameters are entropy (H), anisotropy (A), and alpha (α). They result from the well-known Cloude-Pottier decomposition of the coherency matrix. The potential of Pauli decompositions, which is suitable for coherent targets, has been also been assessed. The segmentation of the data in the Entropy/Alpha (H/α) space has been used to assign each data point to one of 8 classes based on the radar scattering mechanism. The results show that the dominant mechanism from all crops is the surface scattering. This is mainly due to the fact that the scene was acquired during the early growth stage of the crops.

Keywords: Polarimetric SAR, H/α plane, Scattering Mechanism, Wishart Classification.

Nomenclature

LC	Land Cover
SAR	Synthetic Aperture Radar
RADAR	Radio Detection And Ranging
POLSAR	Polarimetric Synthetic Aperture Radar
FP	Full Polarimetric
$H/A/\alpha$	Entropy/ Anisotropy/Alpha
OA	Overall Accuracy
MDA	MacDonald Dettwiler and Associates
H/V	Horizontal/ Vertical
GCP	Ground Control Point
NDVI	Normalized Difference Vegetation Index

1. Introduction

Radio Detection and Ranging (Radar) technology has many advantages. It can penetrate through cloud cover, haze and dust. It has also the capability of day and night imaging. Moreover, radar works with an

active sensor that illuminates the targets by its own energy source. Advanced radar system is the Polarimetric Synthetic Aperture Radar (POLSAR) which can retrieve further information. For these reasons POLSAR data become a useful tool for many applications (i.e. land cover classification). In order to solve the problem of poor classification accuracy, the multichannel measurements should be used. Multichannel measurements can be represented in POLSAR [1]. LC applications (i.e. change detection, crop classification and crop condition monitoring) using optical images are successful only when images can be acquired frequently over the entire crop growth period. However, due to the existence of haze and cloud, high quality optical data are not always available under unfavorable weather conditions. Therefore, when time and area gaps in data acquisition occur, the application potentials of optical images are often limited [2]. In contrast, RADAR remote sensing satellites employ microwaves, which are able to propagate through most cloud and haze. Thus, the backscattering signals obtained are less influenced by weather conditions. The SAR system provides complementary information for optical remote sensing. The backscattering signals from SAR are sensitive to the architecture and dielectric properties of land surfaces, such as plant canopy, built-up areas and soils [2]. [3] Show that the sensitivity of SAR backscattering to crop conditions depends on the SAR sensor parameters (wavelength, incidence angles, and polarization) and also surface parameters (surface roughness, dielectric constant and local incidence angle). Generally, short SAR wavelengths, such as X-band (~3cm) and C-band (~6cm), are less capable to penetrate through the canopy, and therefore mainly interact with the top part of the canopy layers. In contrast, longer wavelengths such as L-band (~20cm) and P-band (~100cm) can penetrate into the vegetation cover and even reach the soil. Early studies mainly focused on the single polarization [4][5] such as ERS-1, ERS-2, JERS-1, and RADARSAT-1, which only provided single or dual polarization data. Single or dual polarization has limited information that cause difficulty in discriminating LC types [6][7][8]. [6] reported that the differential attenuation of the co-polarization



backscatter coefficients or their ratio are useful for grain crops.

The penetration depth achieved depends on the biophysical parameters of the objects causing scatter within a vegetation layer (e.g., water content, size and geometry of the scatter objects), which might enhance or attenuate the interactions between microwaves and scatter-producing features. Full Polarimetric (FP) data can achieved a higher accuracy for LC applications due to much information compared to single or dual polarization [9][2] [10] [11]. The FP data are available through the radar systems, such as C-band RADARSAT-2, X-band TerraSAR-X, and Phased Array type L-band SAR (PALSAR) sensors. [12] showed that the accuracy of Cloude decomposition model and texture analysis is higher than Freeman decomposition in discriminating different types of land covers using neural network method. [13] show that $H/A/\alpha$ decomposition parameters are less effective in crop type classification than Pauli decomposition parameters, but are better than Freeman-Durden decomposition parameters. The overall accuracy (OA) from using $H/A/\alpha$ and Freeman-Durden were 84.1% and 76.9% (OA); respectively. They also demonstrated the potential of the Pauli decomposition parameters for accurate crop mapping in urban/rural fringe areas using Gaussian based Maximum Likelihood classifier. [14] Investigated on the potential of TerraSAR-X, ASAR/ENVISAT and PALSAR/ALOS for monitoring sugarcane crops. Their results show that cross polarizations at long radar wavelengths are mostly sensitive to the changes in sugarcane crops' height and NDVI early in the growing stages.

The objective of this research is to identify the scattering mechanism of each surface in the study area using $H/A/\alpha$ parameters introduced by Cloude-Pottier decomposition. In this context, we retrieved physical information of each surface from the three basic mechanisms: single, double and volume scattering. Moreover, we evaluated the quad RADARSAT-2 data for usage in crop classification in Al-Jimmeza village, Gharbia Governorate, Egypt.

This study has been implemented at system and computer department. Al-Azhar university. The paper is organized as follows: Section (2) focuses on the Study Area and DataSet. Section (3) present the Methodology. Section (4) is about Results and Discussions and section (5) presents the Conclusion.

2. Study Area and DataSet

The study area is Al-Jimmeza village, Gharbia Governorate, located almost in the centre of the Nile Delta, Egypt (30.843 N, and 30.9874 E) (see Figure 1). This site is one of the most agriculturally productive areas in Gharbia Governorate. This area includes the following objects: rice, maize, grape, cotton and urban. The satellite image used RADARSAT-2 image. The RADARSAT-2 satellite was successfully launched on December 14th, 2007

by MacDonald Dettwiler & Associates (MDA) and the Canadian Space Agency, as the successor of the earlier RADARSAT-1. The RADARSAT-2 is a FP one which has four channels of HH, HV, VH, and VV. The "H" indicates horizontal and "V" indicates to vertical transmit or receive. It is worked uses C-band that ranges from 3.5-5.7cm. It is characterized by spatial resolution of 4.5-7.9 m. Due to the small size of the fields in the study area compared to the resolution of the sensor. This affects the boundary of the agriculture fields. Therefore, this should be considered before interpreting and classifying the image.

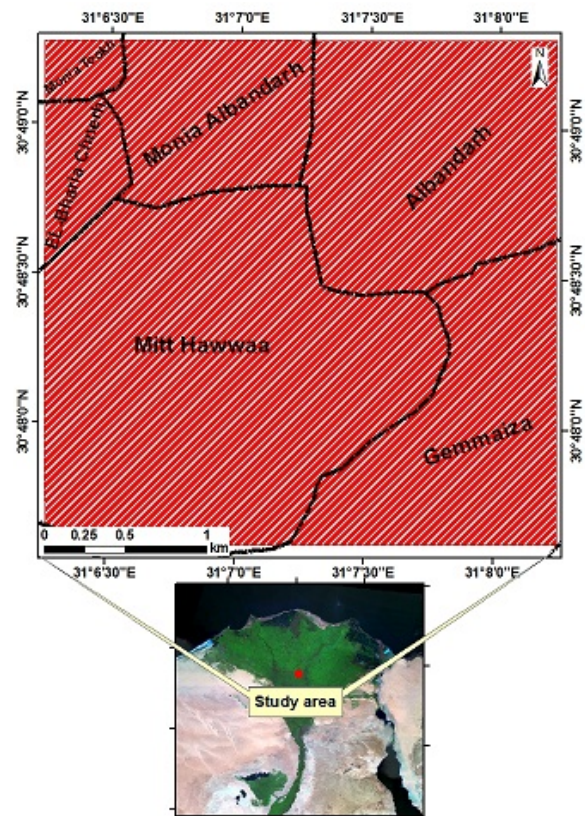


Figure 1 Location map of the study area

3. Methodology

When radar signal interacts with the target the polarimetric SAR measures the four complex components of the backscatter in each resolution cell. This backscatter can be expressed by a scattering matrix (1)

$$\begin{bmatrix} E_h^s \\ E_v^s \end{bmatrix} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \begin{bmatrix} E_h^i \\ E_v^i \end{bmatrix} \quad (1)$$

The superscript "i" refers to the incident wave and "s" refers to the scattered one. The scattering matrix is used to represent the backscatter of the coherent or pure target like urban area. In contrast to the natural target by which partially polarized waves dominate.

To describe the distributed scatterers, the second order matrices are used. These second order matrices are derivatives from the scattering matrix. In case of



reciprocity condition in which $S_{HV} = S_{VH}$ then the vectorized format of the scattering matrix is given in form of lexicographic basis and Pauli basis. The most used one is the Pauli vector which is given by equation (2). Figure 2 shows its composite color of the study area.

$$K_p = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix} \quad (2)$$

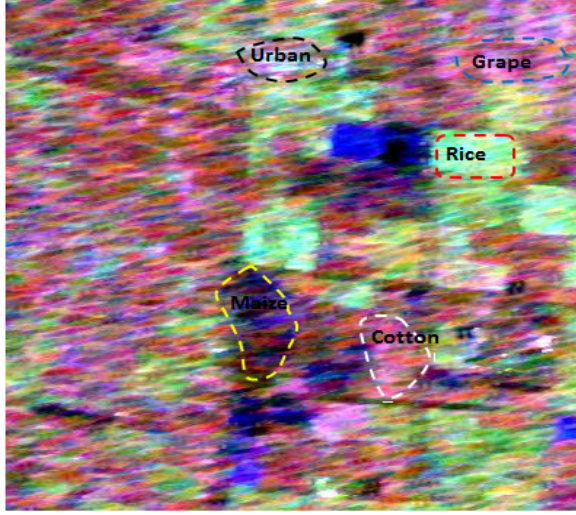


Figure 2 Identification of crop types in Pauli decomposition (RGB)

3.1 Speckle reduction and Geometric Correction

The speckles appear in radar images due to the coherent interference of waves reflected from many elementary scatters [15] and will affect the interpretation of the image. Speckles are maximum in single look images and the spatial resolution is high, in contrast to the multi-look images which the speckles are minimum and the spatial resolution is lower. BoxCar and Refined Lee filters are good at preservation of polarimetric information for distributed targets. However, the Refined Lee filter is better in preserving polarimetric information than boxCar. Refined Lee improves the maintenance polarimetric propertise of the image by considering statistical correlation between channels, i.e. not introducing cross talk [16][17]. RADARSAT-2 data is should be geometrically corrected using ground control points (GCP) when the data are geometrically distorted due to the acquisition system.

3.2 Polarimetric Decomposition

The main objective of polarimetric decomposition is to extract the target information from the backscatter which depends on second order of covariance and coherence matrices, because the first order of scattering matrix is very difficult to use. There are several models for the physical interpretation. However, the polarimetric decomposition is easier and don't require a large number of input parameters compared to others which foreword models are

usually complex and employ large number of input parameters to model the backscatter. There are two types of scatterers: coherent, deterministic or punctual targets are those where the interaction with electromagnetic reflects completely polarized wave. This type of decomposition is pure target like man-made structures. The Pauli decomposition is the most common one of the coherent decomposition. In contrast to the non-deterministic targets or Incoherent targets which are represented by distributed targets and the reflected backscatter measure is the overlapping of a large amount of waves with variable polarization like vegetation areas. $H/A/\alpha$ and Freeman-Durden decompositions are from this type Incoherent decomposition [18]. We are going to perform the non-coherent decomposition proposed by cloude and pottier using the coherency matrix. The cloude decomposition was first presented by cloude in (1996).

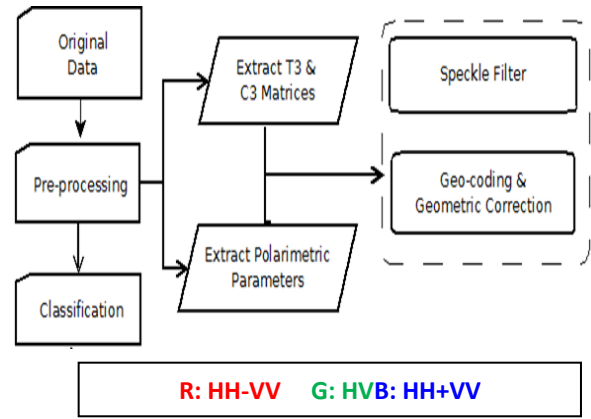


Figure 3 Procedures of processing the Polarimetric SAR image

It is based on the coherency matrix given by (3):

$$[T] = \langle k_p k_p^T \rangle = \sum_{i=1}^3 \lambda_i k_p k_p^T \quad (3)$$

Where the subscript T denotes conjugate transpose, λ_i ($i = 1, 2, 3$), and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$, k_p is the pauli vector (see equation 2)

The diagonal form of the coherency matrix is generated from the eigenvectors and eigenvalues which can be physically interpreted between a set of target vectors. The eigenvalues of $[T]$ therefore have direct physical significance in terms of the components of scattered power into a set of orthogonal unitary scattering mechanisms given by the eigenvectors of $[T]$. Then, the coherency matrix can be written in the form (4):

$$\langle [T] \rangle = [U_3][\Sigma][U_3]^{-1} = \sum_{i=1}^3 \lambda_i u_i u_i^{*T} \quad (4)$$

Where $[\Sigma] = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$ and $[U_3] = [u_1 \ u_2 \ u_3]$

is a unitary matrix, u_1 , u_2 and u_3 are the three unit orthogonal eigenvectors and can be represented in equation (5).



$$[U_3] = \begin{bmatrix} \cos \alpha_1 & \cos \alpha_2 & \cos \alpha_3 \\ \sin \alpha_1 \cos \beta_1 e^{i\delta_1} & \sin \alpha_2 \cos \beta_2 e^{i\delta_2} & \sin \alpha_3 \cos \beta_3 e^{i\delta_3} \\ \sin \alpha_1 \sin \beta_1 e^{i\delta_1} & \sin \alpha_2 \sin \beta_2 e^{i\delta_2} & \sin \alpha_3 \sin \beta_3 e^{i\delta_3} \end{bmatrix} \quad (5)$$

The polarimetric parameters which introduced from the Cloude-Pottier decomposition are based on eigenvalues namely entropy H and anisotropy A and eigenvectors namely α angle. The entropy parameter is used to know the degree of randomness for each target, it ranges from $H=0$ which indicate a single scattering mechanism to $H=1$ which indicate a random scattering mechanism (see Figure 4.a) and is defined from the logarithmic sum of eigenvalues of the coherency matrix as given in equation (6):

$$H = -\sum_{i=1}^3 P_i \log_3(P_i) \quad (6)$$

The alpha parameter α is used to identify the type of scattering mechanism of the target, and it ranges from 0° to 90° degrees. $\alpha = 0$ indicates surface scattering, $\alpha = 45$ indicates volume scattering, and $\alpha = 90$ indicates double bounce scattering (see Figure 4.b). α can be represented using the equation (7)

$$\alpha = \sum_{i=1}^3 P_i \alpha_i \quad (7)$$

Where P_i are the probabilities obtained from the eigenvalues λ_i with

$$P_i = \frac{\lambda_i}{\sum_{j=1}^3 \lambda_j} \quad (8)$$

The anisotropy measures the relative importance of the second and the third eigenvalues of the eigen-decomposition (see Figure 4.c). It is given by:

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \quad (9)$$

4. Results and Discussi

4.1. Entropy-Alpha space

In order to study the type of scattering mechanism for each feature in the study area, H/α plane is used. The scattering mechanism is represented in single scattering, double bounce scattering and volume scattering. This scattering mechanism is along alpha axis and low, medium and high degree of randomness along the entropy axis in H/α plane. The H/α plane

is subdivided into nine zones that characterize each class of different scattering mechanism (see Figure 5.a). Figure 5.b describes the classification of the study area based on H/α space (see Figure 5.c). It is segmented to 8 zones to assign each data point of the study area to one of these 8 zones based on the radar scattering mechanism. The most contribution of this plane is a measure of the number of different types of scattering mechanism found in the averaged samples. When the entropy equal 0, this indicate that there is one dominant scattering as for the entropy equal 1, this indicate that there are all the dominant scattering. The scattering mechanism from each crop can be identified based on the location of the crop's pixels in the H/α space (see Figure 5.a). For example, the rice plant triggers mostly double-bounce scattering based on its high a value. However, its medium entropy suggests that this mechanism is not dominant. Same mechanism applies to the urban scattering. On the other hand, the grape, cotton and maize are located in the medium to high entropy area (i.e. no dominant scattering mechanism) and they occupy a wide range of a value. No unique scattering can be attached to any one of these crops although the grape seem to be relatively well identified with medium α and H . Each zone represents one type of scattering mechanism:

- Z1: High Entropy Multiple Scattering
- Z2: High Entropy Vegetation Scattering
- Z3: (Not a Feasible Region)
- Z4: Medium Entropy Multiple Scattering
- Z5: Medium Entropy Vegetation Scattering
- Z6: Medium Entropy Surface Scattering
- Z7: Low Entropy Multiple Scattering
- Z8: Low Entropy Dipole Scattering
- Z9: Low Entropy Surface Scattering

It is observed also that the dominant scattering with the study area is the surface scattering. Moreover, H/α space has an important contribution in improving the discrimination of rice and urban crops from any other one. The poor classification of maize, grape and cotton crops was occurred due to similar structure with scattering mechanism.



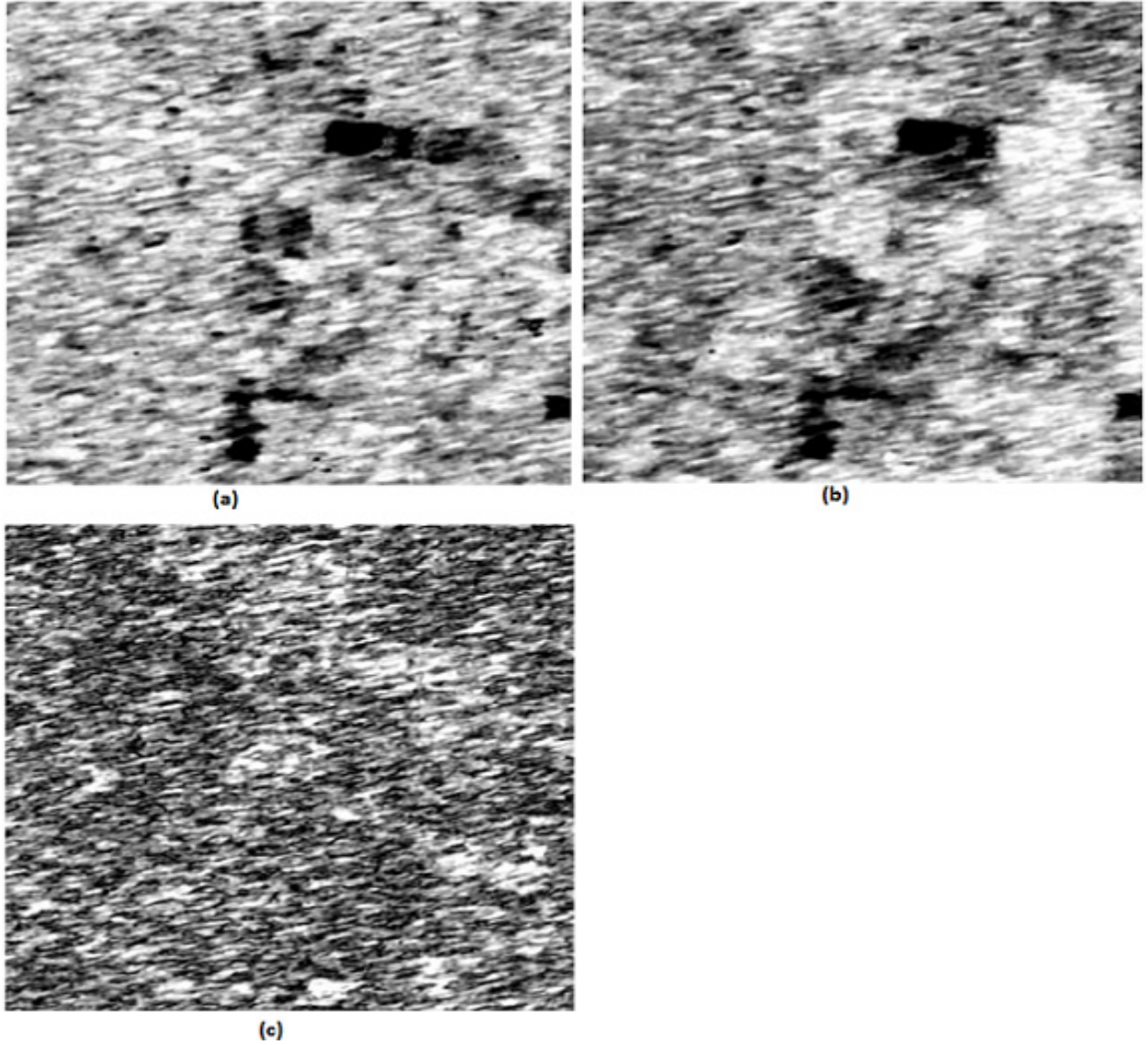


Figure 4 Polarimetric parameters of $H/\alpha/A$ (a) Entropy (b) Alpha (c) Anisotropy

4.2 Wishart Classification

In order to distinguish different classes, the anisotropy parameter is introduced. H/α decomposition has 8 effective zones and by splitting each one into two zones, the number of classes increases from 8 to 16 zones. Therefore, more details can be provided with introducing the anisotropy parameter. H/α and $H/\alpha/A$ decompositions are used as a training sets for the initialization of the unsupervised wishart classification [19][15].

For a coherency matrix T_i of a pixel i of a multilook image (L-looks) knowing the class ω_i , the Wishart complex distribution is given by the following equations[15][20]:

$$p\left(\frac{\langle T_i \rangle}{w_m}\right) = \frac{N^q |\langle T_i \rangle|^{N-q} \exp(-\text{tr}(N [\Sigma_m]^{-1} \langle T_i \rangle))}{K(N, q) |\Sigma_m|^N} \quad (10)$$

Where $\Sigma_m = E(\langle T_i \rangle | \langle T_i \rangle \in w_m)$

$= \frac{1}{N_m} \sum_{i=1}^{N_m} \langle T_i \rangle$ where N_m is the pixel number of w_m and $K(N, q)$ is the factor of standardization given by :

$$K(N, q) = \pi^{\frac{q(q-1)}{2}} \prod_{i=1}^q \Gamma(N - i + 1) \quad (11)$$

Where with reciprocity case ($S_{hv} = S_{vh}$) $q=3$, $\Gamma(\cdot)$ is the gamma function, $| \cdot |$ and $\text{tr}(\cdot)$ indicate the determinant and the trace of the matrix respectively.

A probabilistic measurement of the distance between the matrix of coherence of an unspecified pixel T_i , and the average matrix of coherence Σ_m of the class candidate w_m , is obtained using:

$$d(\langle T_i \rangle, \Sigma_m) = \ln |\Sigma_m| + \text{tr}([\Sigma_m]^{-1} \langle T_i \rangle) \quad (12)$$

Mathematically, each coherency matrix of an individual pixel is assigned with the most likely class ω_m with the minimal distance, if and only if :

$$d(\langle T_i \rangle, \Sigma_m) \leq d(\langle T_i \rangle, \Sigma_n) \quad (13)$$

For all $w_m \neq w_n$



Qualitatively speaking, it appears that the combined H/α parameters and wishart classification provide more details compared to H/α decomposition (i.e. separate the surfaces better). The results are given in Figure 6.a. The crops in the study area is still mixed and are not well delineated their borders. Thus, when using the $H/A/\alpha$ decomposition with the unsupervised wishart classifier, it provides better result (see Figure 6.b). With the second method, we noticed that the contribution of the anisotropy parameter information which is useful for

distinguishing between different cluster centers according to the class identification. The five classes with different degrees of scattering mechanisms can be observed that due to structural parameters. By visual interpretation, if we compare this result with the reference, we can see the difference. This is due to the nature of the unsupervised classification method. Also, this approach of wishart distribution has important factor in discriminating the crops which seem to be better delineated in compared to the RGB composite color

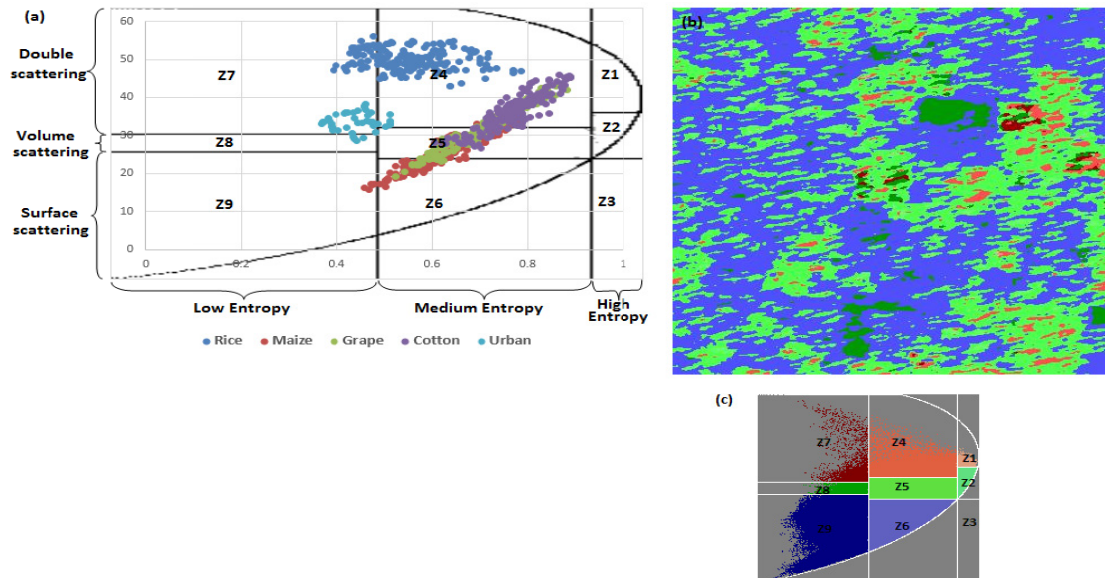


Figure 5 (a) Location of the crop's pixels in the H/A space (b) Segmentation image (c) H/A space

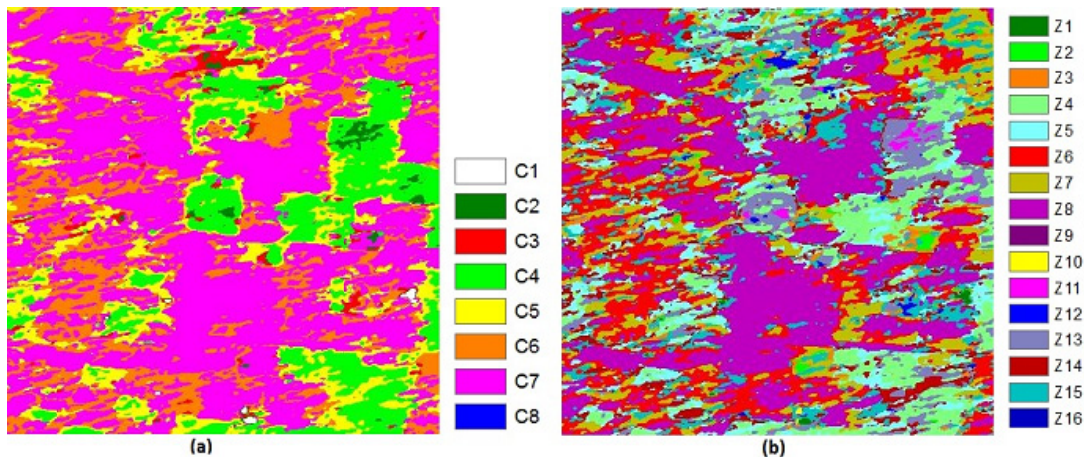


Figure 6 unsupervised wishart classification (a) classification using H/A wishart (b) classification using $H/A/A$ wishart

5. Conclusions

In this paper, we investigate on the unsupervised classification of RADARSAT-2 image that covers Al-Jimmeza village, Gharbia Governorate, Egypt. The classification is represented by two decompositions: H/α and $H/\alpha/A$ based on complex wishart distribution which is the first stage of maximum likelihood. H/α is used to identify the scattering type of each crop. It is observed that the

dominant scattering with the study area is the surface scattering when using H/α plane.

Qualitatively speaking, this classification showed that the different cluster centers could be distinguished using both decompositions. As when using the entropy/anisotropy/alpha decomposition with the unsupervised wishart classifier, it provides better result.



6. Acknowledgment

The authors would like to thank Dr. Sayed Medany and Dr. Mohammed Shokr of NARSS for sharing the Radarsat image used in this study. The authors also would like to thank the Canadian Space Agency for providing the image used in this study within the framework of their Science and Operational Applications Research – Africa (SOAR-AF) under the agreement of the project ID5128. Thanks go also to National Authority for Remote Sensing and Space Sciences, Egypt for offering its resources to perform this study.

7. References

- [1] Y., Maghsoudi, "Analysis of Radarsat-2 full polarimetric data for forest mapping" 2000. Ph.D. Thesis. Reading University: CALGARY,.
- [2] H., McNairn, C., Champagne, Shang, J., D.A., Holmstrom, and G., Reichert, 2009. "Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories" ISPRS Journal of Photogrammetry and Remote Sensing, 64(5), pp. 434-449,
- [3] F. T., Ulaby, A. Tavakoli, and T. B. A. Senior, "Microwave propagation constant for a vegetation canopy with vertical stalks", 1987. IEEE Trans. Geosci. Remote Sens., 25(6), 714–725.
- [4] F., Weber, D., Nixon, J., Hurley, "Semi-automated classification of river ice types on the Peace River using RADARSAT-1 synthetic aperture radar (SAR) imagery", 2003. Canadian Journal of Civil Engineering, vol: 30, pp. 11-27.
- [5] J., Li, W., Chen, R., Touzi, "Optimum "RADARSAT-1 configurations for wetlands discrimination: A case study of the Mer Bleue peat bog", 2007. Canadian Journal of Remote Sensing 33 (SUPPL. 1), pp. S46-S55
- [6] F. T., Ulaby, F., Kouyate, B., Brisco, and T. H. L. Williams, "Textural information in SAR images", 1986. IEEE Transactions on Geoscience and Remote Sensing, 24, 235–245,
- [7] X., Li, A. G. O., Yeh, J. P., Qian, B., Ai, and Z. X., Qi, "A matching algorithm for detecting land use changes using case-based reasoning", 2009. Photogrammetric Engineering and Remote Sensing, 75, 1319–133.
- [8] D., Wang, H., Lin, J., Chen, Y., Zhang, and Q., Zeng, "Application of multi-temporal ENVISAT ASAR data to agricultural area mapping in the Pearl River Delta, International Journal of Remote Sensing", 2010. IEEE Journal of Selected Topics in Applied Earth Observations and Remote, vol: 31, pp.1555-1572.
- [9] J.S., Lee, and E., Pottier, "Polarimetric radar imaging: from basics to applications. 2009. Taylor & Francis Group, New York.
- [10] X., Niu, and Y., Ban, "Multi-temporal RADARSAT-2 polarimetric SAR data for urban land-cover classification using an object-based support vector machine and a rule-based approach", 2013. International Journal of Remote Sensing, 34(1), 1-26
- [11] Z., Qi, A. G. O., Yeh, X., Li, and Z., Lin, "A novel algorithm for land use and land cover classification using RADARSAT-2 polarimetric SAR data. 2012. Remote Sensing of Environment", 118, 21-39.
- [12] K., Sambodo, A., A., Murni, and M., Kartasasmita (2007), "Classification of polarimetric-SAR data with [19] K., Sambodo, A., A., Murni, and M., Kartasasmita (2007), "Classification of polarimetric-SAR data with neural network using combined features from scattering models and textural analysis", Remote sensing and earth sciences, Vol. 4.
- [13] Q., Ma, J., Wang, J., Shang and P., Wang 2012. "Assessment of Multi-temporal RADARSAT-2 Polarimetric SAR Data for Crop Classification in an Urban/Rural Fringe Area", IEEE international (IGRASS). 314 – 319
- [14] N., Baghdadi, N., Boyer, P., Todoroff, M., El Hajj, A., Bégué (2009), "Potential of SAR sensors TerraSAR-X, ASAR/ENVISAT and PALSAR/ALOS for monitoring sugarcane crops on Reunion Island" Remote Sensing of Environment, Vol.113, pp.1724-1738.
- [15] J. S., Lee, M. R, Grunes, T. L., Ainsworth, L. J., Du, D. L., Schuler and S. R., Cloude, "Unsupervised classification using polarimetric decomposition and the complex Wishart distribution," 1999. IEEE Transactions Geoscience and Remote Sensing, Vol 37/1, p 2249-2259,.
- [16] M., Salehi, M. R., Sahebi, and Y., Maghsoudi, "Improving the accuracy of urban land cover classification using radarsat-2 polSAR data", 2013. IEEE journal in selected topics in applied earth observations and remote sensing.7: 1394 – 1401.
- [17] C., Niu, G., Zhang, J., Zhu, S., Liu, and D., Ma, "Correlation coefficients between polarization signatures for evaluating polarimetric information preservation," 2011 IEEE Geosci. Remote Sens. Lett, 8 :1016–1020.
- [18] R., Verma, "Polarimetric decomposition based on general characterisation of scattering from urban areas and multiple component scattering model", 2012. M.Sc. Thesis. Reading University: Twente
- [19] S. R., Cloude, and E., Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," 1997. IEEE TGRS, vol.35, pp. 68-78,
- [20] M., Ouarzeddine, B., Souissi, A., Belhadj-Aissa, 2012. "Classification of polarimetric SAR images based on scattering mechanisms", ISPRS, vol: XXXVI.



Biography



Hind H. Zeyada is an engineer in National Authority for Remote Sensing and Space Sciences. She is currently data analyzer at the Data Reception, Analysis and Receiving Station Affairs Division. She received her B.Sc. in Computers and Systems Engineering Department, Al-Azhar University. Her area of interest includes image processing and GIS in different fields of Remote Sensing. Mail: sar.hend87@gmail.com



Mohamed M. Ezz is Assistant Professor in Faculty of Engineering Al Azhar University. He received his B.Sc., M.Sc. and Ph.D. in Systems & Computers Engineering from Faculty of Engineering, Al Azhar University. He is IEEE member. His area of interest includes application security, intrusion detection, semantic web, and pattern recognition. He has published 20 scientific papers in various national and international journals and conferences. He has contributed in more than 16 mega software projects in Electronic banking EBPP, EMV, mobile banking and e-commerce. Mail: Ezz.mohamed@gmail.com



Ayman H. Nasr is Associate Research Professor in National Authority for Remote Sensing and Space Sciences. He is currently the Head of the Data Reception, Analysis and Receiving Station Affairs Division. He received his B.Sc., M.Sc. and Ph.D. in Electronics and Communications from Faculty of Engineering, Cairo University. His area of interest includes image processing and GIS in different fields of Remote Sensing. He has published 30 scientific papers in various national and international journals and conferences. He has contributed in publishing 4 atlases and participated in more than 48 projects and researches related to remote sensing fields. Mail: aymanasr@hotmail.com



Hany M. Harb is a professor in Faculty of Engineering at Al-Azhar University. He is currently the Head of the Computers and Systems Engineering Department. He received his Doctor of philosophy (Ph.D.), Computer Science, Illinois Institute of Technology (IIT), Chicago , Illinois , USA, 1986. He received his Master of Science in Operations Research (MSOR), IIT, 1987. He received his Master of Science in Computers and Systems Engineering, Faculty of Engineering, Azhar University , Cairo , Egypt, 1981. He received his Bachelor of Science in Computers and Control Engineering, Faculty of Engineering, Ain Shams University, Cairo, Egypt, 1978. His area of interest includes object Oriented Software Engineering Modeling Using UML, Software engineering, Distributed Systems and Computer Networks. Mail: harbhany@yahoo.com





Hamming Graph based Image Compression with variable Threshold

Levstein F.¹, Lezama J.², Maldonado C.³, Penazzi, D.⁴, Schilman M.⁵

Universidad Nacional de Córdoba: ^{1,2,4,5}: FaMAF. ³: FCEFyN. ^{1,2,3}: CONICET.

levstein@famaf.unc.edu.ar, javitolez@gmail.com, cmaldona@gmail.com, danielpenazzi@gmail.com,
maurito.s@gmail.com

<http://www.famaf.unc.edu.ar>, <http://www.portal.efn.uncor.edu>

Abstract

In this paper a lossy image compression method using the adjacency matrix of a graph is proposed. We instantiate the method using the Hamming graph, which gives rise to an image compression algorithm that uses a Hadamard Matrix together with a threshold approach. In threshold methods, selection of the correct (static) threshold value is difficult. In our approach we use a new threshold determination that changes from block to block obtaining an algorithm that uses only integer and modular operations and obtains good results in some standard images. The new method has better stability under recompression iteration than jpg.

Keywords: Image compression; Hadamard transform; Threshold approach.

1 Introduction

One of the techniques used in image compression is to decompose the vector space associated to subblocks of the image into orthogonal subspaces, and keep some of the projections instead of the total image. ([1, 7, 13, 14, 15, 20, 24]).

In this work we consider a spectral decomposition of \mathbb{R}^{64} associated with the eigenspaces of the adjacency matrix A of a Hamming graph ([1, 4]). The Hamming graph gives rise to a compression algorithm that uses a Hadamard matrix. This has been used extensively before (e.g. [1, 2, 7, 9, 16, 17, 19, 25, 29, 31]). Our work differs in which projections are discarded, and in the determination of the threshold that determines into how many eigenspaces we project.

We first (section 2) exemplify the general idea, recalling the definition of Hamming graph and exemplifying with a basic image compression algorithm in which we divide the image in blocks of size 2×2 pixels. In section 3 we begin the main contribution of the paper, introducing the basics of the main algorithm, which involves blocks of size 8×8 . In this section we

also introduce the variable block threshold approach we use and show some results and discuss some shortcomings of the algorithm. Then in section 4 we show some tweaks done to the algorithm to improve both the level of compression and the quality of the visual results. In subsection 4.2 we discuss the results of testing the complete algorithm on a variety of images from [33] and in section 4.3 we discuss the stability under recompression iteration of our algorithm, comparing it favorably against jpeg. In the appendix we show the complete table of experiments.

2 Overview

This section reviews basic definitions of the objects we will use.

2.1 General Idea

Let A be a symmetric matrix of order $m \times m$, for example the adjacency matrix of a graph with m vertices. Recall that the adjacency matrix A of a graph G is the matrix indexed by the set of vertices X and defined by:

$$(A)_{x,y} = \begin{cases} 1 & \text{if } (x,y) \text{ are adjacent in } G \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

Since we are going to be interested in applications to images with square subblocks we will take $m = n^2$ for some n . Such a symmetric matrix gives an orthogonal decomposition of \mathbb{R}^m into eigenspaces of A ([3, 4, 5]). Thus there exists an orthogonal basis $B = \{v_1, v_2, \dots, v_m\}$ of \mathbb{R}^m consisting of eigenvectors of A .

A basic compression method consists on keeping only the coefficients of some of the projections to the eigenspaces. If the eigenspaces are ordered according to the magnitude of the eigenvectors, eliminating the projections of the lower eigenvalues should provide a good reconstructed image.



In this article we will be using as our graphs the Hamming graphs. We next review them briefly.

2.2 Hamming Graphs

Definition 2.1. Let $X = \mathbb{Z}_2^d$. For x and y in \mathbb{Z}_2^d , x, y are adjacent if they differ in exactly one position, that is $|\{i : x_i \neq y_i\}| = 1$. The graph $H(d, 2)$ consist on the set of 2^d vertices X and the adjacency relation previously defined (see [4, 5]).

Example 2.2. Example: $H(2, 2)$:

For $d = 2$ the set of vertices of the graph is

$$X = \mathbb{Z}_2^2 = \{(0, 0); (0, 1); (1, 0); (1, 1)\} \quad (2.2)$$

and with that order of the vertices the adjacency matrix A is

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad (2.3)$$

Example 2.3. Example: $H(3, 2)$:

For $d = 3$ the set of vertices of the graph is $X = \mathbb{Z}_2^3 = \{(0, 0, 0); (0, 0, 1); (0, 1, 0); (0, 1, 1); (1, 0, 0); (1, 0, 1); (1, 1, 0); (1, 1, 1)\}$ and with that ordering of the vertices, the adjacency matrix A is

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \quad (2.4)$$

2.3 Expository Case: $d = 2$

In the next section we begin to explain our algorithm, which is based on the case $d = 6$, but for clarity of exposition we give next a very simple example which shows how to compress, using an orthogonal basis obtained from the spectral decomposition of \mathbb{R}^4 with $H(2, 2)$.

For $d = 2$ the eigenvalues, ordered from highest to lowest, are $\lambda_0 = 2$, $\lambda_1 = 0$ and $\lambda_2 = -2$ and we obtain the decomposition:

$$\mathbb{R}^4 = V_{\lambda_0} \oplus V_{\lambda_1} \oplus V_{\lambda_2} \quad (2.5)$$

with:

$$\begin{aligned} V_{\lambda_0} &= \langle (1, 1, 1, 1) \rangle \\ V_{\lambda_1} &= \left\langle \begin{pmatrix} 1, -1, 1, -1 \\ 1, 1, -1, -1 \end{pmatrix} \right\rangle \\ V_{\lambda_2} &= \langle (1, -1, -1, 1) \rangle \end{aligned}$$

Arranging the eigenvectors on a matrix we obtain:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad (2.6)$$

that is, a Hadamard matrix.

Recall that a Hadamard matrix is any matrix with entries ± 1 in which the rows are orthogonal to each other. A Hadamard matrix H of order n satisfy $H^t H = H H^t = nI$. Hadamard matrices have been used for quite some time in image compression algorithms, for example, in [1] the authors show the advantage of using Hadamard due to its computational speed; in [29] there is further discussion of the advantages of using the Hadamard matrix by showing an algorithm of image compression that uses this matrix together with an adaptative clustering technique. The algorithm compares favorably with jpeg and others. There are also applications in biomedicine focusing on a particular class of medical images in grayscale: magnetic resonance imaging (MRI) of the human brain. In [2] the authors focus on document image compression. They propose a new transform, which they call rCS-SCHT and is derived from a Hadamard transform. Like the Hadamard transform, the rCS-SCHT requires only addition and subtractions. The authors do tests of the new transform and compare it to the discrete cosine transform and the Walsh Hadamard transform in relation with scanned document images. In [25] the authors propose the use of a one-dimensional Hadamard naturalness preserving transform (NPT), considering the problem of reconstructing a unidimensional time variable signal from segments of its Hadamard NPT and an iterative Hadamard NPT reconstruction algorithm is introduced; in [31] is like the previous but using a two dimensional NPT, in [7] the Hadamard matrix is used together with a new quantization matrix based on the Human Visual System; In [17] the structure of the Walsh matrices and Hadamard matrices are briefly discussed. In [11] a Hadamard transform (together with a Discrete Wavelet Transform in a cascaded transformation) is used on a watermarking scheme for color images. [19] is one of the many applications to video compression, using a Variable Block Size (VBS) partition algorithm that utilizes the Walsh Hadamard Transform. [9] also use the Discrete Walsh-Hadamard Transform in a video coding scheme. They show that for low bitrate applications the performance is analogous to those obtained with the Discrete Cosine Transform in terms of compression efficiency, PSNR, and visual quality, but with ar less computational complexity of the Hadamard transform. [30] uses a single-scan 2D Hadamard approach in the area of two-dimensional nuclear magnetic resonance. Hadamard matrices have even some applications in cryptography (actually, they are "pseudo Hadamard" transformation, since the multiplications by -1 are replaced



by multiplications by 2), for example in the SAFER family of ciphers ([21, 22, 23]).

In [27] both SAFER+ and the U.S. standard AES are implemented in VHDL, with much better throughput for SAFER+. The authors conclude that SAFER+ is better for implementation on Bluetooth devices. (although there is no discussion of the relative security of the algorithms).

Other examples of use of Hadamard Matrices in image compression or image processing are [8, 16, 18, 28]

These are just a few of the many examples that use Hadamard matrices. The advantage of using Hadamard matrices over some other transform, like the Discrete Cosine Transform, is precisely that all entries are ± 1 , so only addition and subtraction operations are required, no multiplications. (A division by n is required either in the compression or decompression, since by the property $H^t H = H H^t = nI$ then $H^{-1} = \frac{1}{n} H^t$. However, since in general n will be a power of 2, this division is implemented quite efficient in computer processors, since an (integer) division by 2^r is the same as a right shift by r bits). We showed above as a way of example that $H(2, 2)$ gives rise to a Hadamard matrix. The same is true for all Hamming graphs. (see for example [10, 26, 5]). If the rows are ordered correctly, the matrix obtained is symmetric, thus $H^{-1} = \frac{1}{n} H$.

The compression method in the case $H(2, 2)$ consists in eliminating the coordinate corresponding to V_{λ_2} . This gives a compression ratio of 4:3 so it is not very good, but as we previously said we include it for expository purposes.

Of course we could try to eliminate also the coordinates corresponding to V_{λ_1} , obtaining a better ratio, but then the reconstructed image bears little resemblance to the original.

Therefore, the compression consists of mapping $v = (a, b, c, d)$ into:

$$w = \frac{1}{4}(a+b+c+d, a-b+c-d, a+b-c-d) \quad (2.7)$$

and the decompression consists of mapping $w = (\alpha, \beta, \gamma)$ into:

$$\tilde{v} = (\alpha + \beta + \gamma, \alpha - \beta + \gamma, \alpha + \beta - \gamma, \alpha - \beta - \gamma) \quad (2.8)$$

The expectation being that v and \tilde{v} would look similar enough to the human eye. (other alternatives would be to not divide by 4 when compressing, dividing instead when decompressing, or, more symmetrical, dividing by 2 both on compression and decompression. Highest compression is obtained by dividing when compressing).

3 The Initial Algorithm

3.1 Overview of the basic Algorithm

As we said previously, our main focus in this article will be the case $d = 6$. In the cases of $d > 2$ the main difference with $d = 2$ is that we can eliminate the coordinates of more eigenspaces, obtaining better compression ratios while still getting good reconstruction.

The case $d = 6$ is equivalent to a multiplication by a Hadamard matrix of size 64×64 , and eliminating some of the coordinates. The problem consists on exactly how many eigenspaces to eliminate, that is where to fix the threshold. In this paper we use a dynamic approach to the selection of the threshold, that is, instead of fixing beforehand the same threshold for every block, we change the threshold from block to block. We explain the details in the next section, together with a discussion of the shortcomings. In a later section, we show a more advanced algorithm to solve this problems.

3.2 Choosing the eigenspaces to be eliminated

For $d = 6$ we have:

$$\mathbb{R}^{64} = V_{\lambda_0} \oplus V_{\lambda_1} \oplus V_{\lambda_2} \oplus V_{\lambda_3} \oplus V_{\lambda_4} \oplus V_{\lambda_5} \oplus V_{\lambda_6} \quad (3.1)$$

with $\lambda_j = 6 - 2j$, i.e., $\lambda_0 > \lambda_1 > \dots > \lambda_6$.

Eliminating only $V_{\lambda_6} = V_{-6}$ would give excellent visual results but poor compression, while eliminating more eigenspaces would give progressively worse visual results. Notice also that the dimensions are $\dim(V_{6-2j}) = \binom{6}{j} \cdot ([6])$

Fixing the number of eigenspaces to be eliminated produce results that are not uniformly good on images, so we let the program decide by itself how many eigenspaces need to be eliminated. This decision is made block by block, so on some blocks almost all eigenspaces are eliminated, while on others very few. This also depends on the image, so the level of compression cannot be predicted beforehand.

The procedure to decide how many eigenspaces to eliminate on a given block is the following:

For each $j > 0$ let π_j be the projection to $\Lambda_j = V_{\lambda_0} \oplus V_{\lambda_1} \oplus \dots \oplus V_{\lambda_j}$. Given an 8×8 block of pixels we take the 64-byte vector that represents the block by reading it column by column. We then check if:

$$\|\pi_j w\|^2 > c \|w\|^2 \quad (3.2)$$

where c is some constant close to 1.

We keep the first $j \geq 1$, (the *threshold*), for which the test is satisfied. The reason we only allow $j \geq 1$ is that tests showed that if $j = 0$ is allowed, the results are very poor visually. (Note: in the finished



algorithm we allowed for all $j \geq 1$, but in the initial algorithm we only allowed thresholds $j=1,2,3,6$ to speed up the tests).

3.3 Results and Discussion of Experiments with initial algorithm.

We chose some images commonly used in the testing of compression algorithms, based on their contrast or intensity, boundary conditions, etc.

We tested our algorithm with several constants c close to 1. Fixing a c , then for each 8×8 block, we calculated the *threshold* j described above.

We focused on three of the constants which provide good visual and metrics results. The constants we focused on were $c = 0.96, c = 0.99, c = 0.999$.

Below there is a table that shows the compression rate obtained for different images and different values of c . The data in red correspond to that images which have good ratio compression as well as good values for the metrics and visual results with the correct explained above constant c :

c	0,999	0,99	0,96
Lena	5:2	5 : 1	8:1
Baboon	7:6	2:1	4 : 1
Blackboard	4 : 1	7:1	8:1
PC	6:2	5 : 1	8:1
Eclipse	8 : 1	9:1	9:1

Example 3.1. Visual results obtained with the image of Lena with $c = 0.99$:



Figure 1: Orig.



Figure 2: 5:1

Example 3.2. Visual results obtained with the image of Baboon with $c = 0.96$

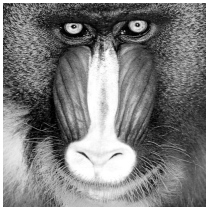


Figure 3: Orig.

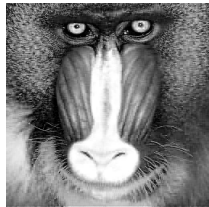


Figure 4: 4:1

For purposes of studying the properties of the c 's we registered for each j the number of blocks projected to Λ_j and analyzed the distribution of the blocks to each Λ_j . Good visual and metrics results were obtained in those cases in which the distribution

of the number of blocks to Λ_1 corresponded to between 60 and 70 percent of the total, the number of blocks to Λ_2 were less than a third of the ones distributed on Λ_1 and the number of blocks in Λ_6 did not exceed 100.

Example 3.3. The distribution of the blocks of Lena with $c = 0.99$ to the corresponding projections Λ_j are the following:

	Λ_1	Λ_2	Λ_3	Λ_6
Lena	3180	520	313	83

In this case it is obtained a compression of $4096 \times 64 = 262,144$ of the originals bytes to: $3180 \times 7 + 520 \times 22 + 313 \times 42 + 83 \times 64 = 52158$ bytes. ($\frac{262144}{52158} = 5.03$ approximately 5 to 1).

Example 3.4. The distribution of the blocks of Baboon with $c = 0.96$ to the corresponding projections Λ_j are the following:

	Λ_1	Λ_2	Λ_3	Λ_6
Baboon	2691	769	565	71

In this case we obtain a compression of $4096 \times 64 = 262,144$ of the originals bytes to: $2691 \times 7 + 769 \times 22 + 565 \times 42 + 71 \times 64 = 64029$ bytes ($\frac{262144}{64029} = 4.09$ approximately 4 to 1).

3.4 Some Tests

In what follows we list the results of the different metrics corresponding to the different c 's. Let's recall briefly the definitions:

The Normalized Cross correlation (NCC):

$$NCC(f, g) = \frac{\sum (f(x) - \bar{f})(g(x) - \bar{g})}{\sqrt{[\sum (f(x) - \bar{f})^2] [\sum (g(x) - \bar{g})^2]}} \quad (3.3)$$

where the sum is over all $x \in X$ and \bar{f} and \bar{g} are the averages of the original and obtained image respectively.

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=1}^{N-1} \|I(i, j) - K(i, j)\|^2 \quad (3.4)$$

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (3.5)$$

(here MAX_I is the maximum value that a pixel can take in an image). The results are as follows:

$c = 0.999$	PSNR	NCC
Lena	41.2989	0.998941
Baboon	42.2399	0.999374
Blackboard	38.4096	0.996655
PC	43.0004	0.999824
Eclipse	57.7279	0.999944



$c = 0.99$	PSNR	NCC
Lena	30.7488	0.999331
Baboon	29.9242	0.989220
Blackboard	34.5211	0.992961
PC	33.4597	0.998386
Eclipse	43.7904	0.998595

$c = 0.96$	PSNR	NCC
Lena	30.3827	0.986472
Baboon	29.1284	0.967854
Blackboard	30.8914	0.983637
PC	28.8779	0.995350
Eclipse	39.2106	0.995959

Notice that with $c = 0.999$ the PSNR and NCC are quite good, however in most cases the compression ratio we obtain is not what we wanted. With $c = 0.99$ and 0.96 the ratios improve but now the PSNR and others are not as good.

This version has a number of other problems that included some artifacts. Therefore, we moved on to a better version.

4 Advanced Algorithm

4.1 Improvements

As we explained the version detailed above has a number of problems that includes some artifacts on certain blocks, that the compression was not good enough and that the statistics PSNR, MSE, etc were not as good as we wanted, and that the packing of the compressed image was not efficient.

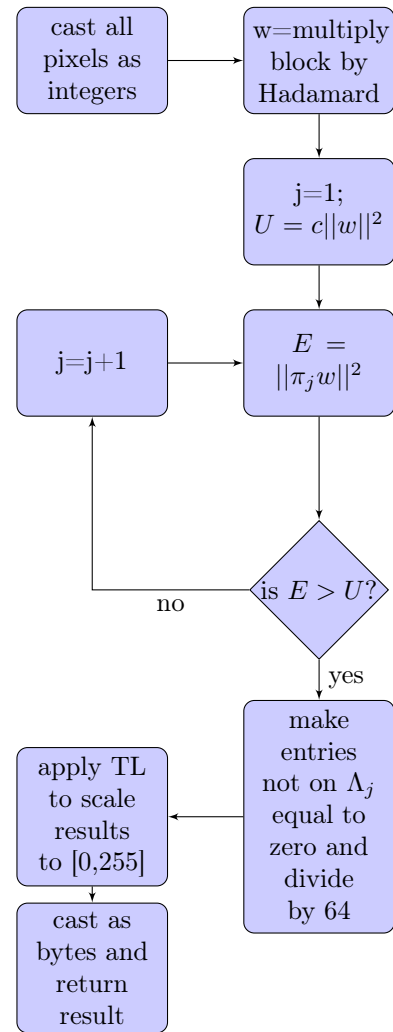
First, we tackled the problem of the efficiency of the packing of the compressed image. Thus, instead of simply eliminating the entries, we made them zero, and then used a lossless file compression algorithm to further compress the resulting image. This gives us much better results than the previous approach.

Another problem was with those pixels with values near 0 or 255. The theory behind the projection and its inverse assumes we are working on the reals, or at least the rationals, but since we are working on pixels=bytes=integers modulo 256, and moreover we are dropping some coordinates and then reconstructing the image, values near (but below) 255 may end up after reconstruction near 255 but above, which will be rendered as near 0. So a white pixel may end up black and viceversa.

To solve this problem we cast the pixels as integers instead of bytes (actually they are integers between -2^{31} and 2^{31} , but for our purposes they behave like integers), and operate as integers.

We make the coordinate change multiplying by the Hadamard matrix (which, again, has all entries ± 1) and then dividing by 64. (the division is integer division, that is, the result is the floor of the division by 64).

Because each original coordinate was between 0 and 255, and the entries of the Hadamard matrix are all 1 or -1 and each row has 64 entries, then it follows that the new coordinates will be between -255 and 255. Thus, we do a scaling: to these coordinates we apply a linear transformation $TL(x) = (x + 256)/2$ so the new coordinates are between 0 and 255. We recast these new entries as bytes, make the entries corresponding to the coordinates that are going to be eliminated 0, and this is the (partially compressed) image returned, to which we apply the lossless compression algorithm. Below we show the flow chart for each block. After all blocks are processed, the lossless compression algorithm is applied to the whole file.

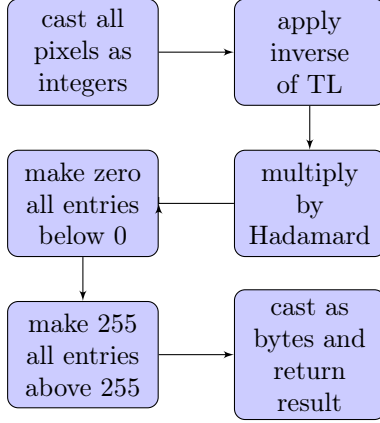


Flowchart 1

To decompress the image, first we decompress using the lossless decompression algorithm, then we apply to each pixel (casted again as integers) the inverse of TL , namely $z \mapsto 2z - 256$, and then we multiply everything by H . Since the original values were between 0 and 255 and $\frac{1}{64}H^2 = Id$, then the values that we do obtain should also be between 0 and 255, except for some round off errors. Thus, as a final step (we are still seeing everything as integers) we send to 0 all those



values below 0, and to 255 all those values above 255. We then recast everything as bytes (pixels) and that is the image returned.



flowchart 2

4.2 Results

We tested the complete algorithm on a variety of standard black and white images taken from [33].

For the lossless compression algorithm incorporated in the design, we tested both gzip and bzip2. We also tested as a way of comparison the results of the algorithm with $d = 6$ with similar algorithms (also using the linear transformation TL and the lossless compression) with $d = 2$ and $d = 4$. In the case of $d = 6$, the constant c used was $c = 0.99$. We also compared with the results obtained by losslessly compressing the image using only gzip or bzip2 (with no algorithm of our own) to test what is the contribution of our algorithm.

We compressed three small images of size 65KB, obtaining sizes between 14 to 27KB with our best algorithm, with PSNR between 29 and 33dB. We compressed 2 images of size 406KB obtaining 45KB with PSNR 34dB and 28KB with PSNR 36dB. We compressed four images of size 1000KB obtaining images of size between 23 and 190KB, with PSNR between 32 and 41dB. We compressed eighteen images of size 257KB obtaining images of size between 22 and 78KB, with PSNR between 30 and 36dB.

The experiments showed that indeed $d = 6$ compress more than the $d = 2$ or $d = 4$ versions and much more than just using gzip or bzip2, while the quality of the reconstruction with $d = 6$ is comparable (and in some cases even better) than the $d = 2$ or $d = 4$ cases. In an appendix we show all the results with the bzip2 model, since it was the one which gave better results. (for reasons of space we do not include the $d = 2$ or $d = 4$ models).

4.3 Iterative Recompression: Comparison with .jpg

We modified the algorithm to be able to process color images, simply by applying the above algorithm to each of the three color bands. (we did not use YC_bC_r encoding).

We then compared the level of degradation when iterating the algorithm, that is, we did the following: call level 0 the original image. We then compress the image, decompress it, and save the decompressed image as level 1. We then repeat the above with the level 1 image, obtaining a level 2 image, etc. We want to compare the differences between images of consecutive levels, to see how much extra degradation is incurred when iterating the procedure. We obtain the following results for Lena: ($d = 6$ compresses a colored Lena from 768KB to 80KB, and the jpg settings were set to compress to 89KB):

Let's recall that the structural similarity (SSIM) index is:

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (4.1)$$

with

μ_x the average of x ;

μ_y the average of y ;

σ_x^2 the variance of x ;

σ_y^2 the variance of y ;

σ_{xy} the covariance of x and y ;

$c_1 = (k_1L)^2, c_2 = (k_2L)^2$ two variables to stabilize the division with weak denominator;

L the dynamic range of the pixel-values (typically this is $2^{\# \text{ bits per pixel}} - 1$;

$k_1 = 0.01$ and $k_2 = 0.03$ by default.

Also, we use:

$$MS_SSIM(X, Y) = \frac{1}{M} \sum SSIM(x_j, y_j) \quad (4.2)$$

where X and Y are the reference and the distorted images, respectively; x_j and y_j are the image contents at the j th local window; and M is the number of local windows of the image.

	Level 1 vs Level 0	
	$d = 6$	jpg
MSE :	26.5	24.1
PSNR :	33.89dB	34.31dB
SSIM :	0.98033	0.99237
MS_SSIM :	0.96990	0.99124

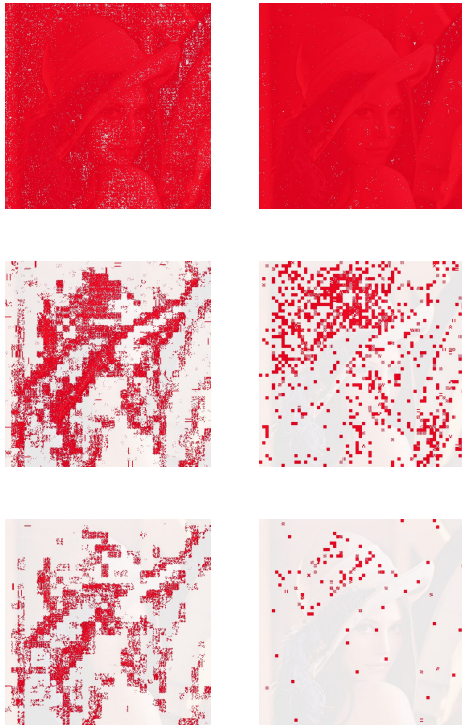
	Level 2 vs Level 1	
	$d = 6$	jpg
MSE :	0.7	4.3
PSNR :	49.49dB	41.77dB
SSIM :	0.99970	0.99933
MS_SSIM :	0.99962	0.99887



	Level 3 vs Level 2	
	$d = 6$	<i>jpg</i>
<i>MSE</i> :	0.1	0.8
<i>PSNR</i> :	57.14dB	49.05dB
<i>SSIM</i> :	0.99996	0.99983
<i>MS_SSIM</i> :	0.99995	0.99975

	Level 4 vs Level 3	
	$d = 6$	<i>jpg</i>
<i>MSE</i> :	0.0	0.4
<i>PSNR</i> :	66.19dB	51.61dB
<i>SSIM</i> :	1.00000	0.99988
<i>MS_SSIM</i> :	1.00000	0.99982

In the figures, each (red in the original) dot indicates a difference between the compared figures, thus the fewer dots the better. The left ones are jpg, the right ones ours.



We see that our algorithm is comparable with jpg in the first iteration, but after that it always outperform jpg, and essentially becomes a fixed point after the fourth iteration. We tried to see when this would happen with .jpg but even after two more iterations this did not happen:

	jpg: Level 5 vs Level 4	
<i>MSE</i> :	0.2	
<i>PSNR</i> :	54.26dB	
<i>SSIM</i> :	0.99994	
<i>MS_SSIM</i> :	0.99992	

	jpg: Level 6 vs Level 5	
<i>MSE</i> :	0.2	
<i>PSNR</i> :	56.06dB	
<i>SSIM</i> :	0.99997	
<i>MS_SSIM</i> :	0.99996	



5 Conclusions and Future Work

We introduced a new algorithm for image compression based on the projections on the eigenspaces of the adjacency matrix of the Hamming graph, with a new method of block-variable threshold determination. The algorithm has better stability under iteration compared to the standard jpg and improves the performance of [25] previously described, where the compression iteration have relevant importance in medical images. We tested the algorithm in several images from [33] obtaining good results. Although the compression step has to compute the optimum threshold, this step is not needed for decompression, so decompression, which is the critical step in bulk image retrieval, is much faster.

We are working with different graphs to obtain better results, particularly the Johnson graph. We are also seeking to incorporate other techniques like overlapping, Huffman code encoding, the human visual system (HVS), etc. to improve the results and have new algorithms for image compression.

References

- [1] Hadamard Transform Image Coding. Andrews Harry C., Kane Julius and Pratt William K.. Proceedings of the IEEE, Vol. 57, NO. 1, January 1969.
- [2] A new transform for document image compression. A. Aung, Boon Poh Ng, C. T. Shwe, in Proc. of 7th International Conference on Information, Communications and Signal Processing, 2009.
- [3] Linear Algebra Done Right. Sheldon Axler Springer Verlag 1997
- [4] Association Schemes: Designed Experiments, Algebra and Combinatorics. R. A. Bailey. Cambridge studies in advanced mathematics 84. Cambridge University Press. 2004.
- [5] Algebraic Combinatorics I: Association Schemes. E. Bannai, T. Ito. Benjamin Cummings, Lecture Notes Series, Menlo Park, Cal, 1984.
- [6] Brouwer, A. E.; Cohen, A.; Neumaier, A. Distance-regular graphs.. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge,



18. Berlin etc.: Springer-Verlag. xvii, 495 p. (1989).
- [7] Designing Quantization Table for Hadamard Transform based on Human Visual System for Image Compression. B. N. Chatterji, S. Srinivaskumar, K. Veeraswamy. ICGST-GVIP Journal, Volume 7, Issue 3, November 2007.
- [8] Application Of Hadamard, Haar, And Hadamard-Haar Transformation To Image Coding And Bandwidth Compression. Ryszard S. Choras; Proc. SPIE 0359, Applications of Digital Image Processing IV, 336 (March 17, 1983); doi:10.1117/12.965983.
- [9] A low complexity video coder based on discrete Walsh Hadamard transform. R. Costantini, J. Bracamonte, G. Ramponi, J. Nagel, M. Ansonge and F. Pellandini, , In EUPSICO 2000 : European signal processing conference, pages 1217-1220, 2000.
- [10] Eigenspaces of graphs. D. Cvetković, P. Rowlinson and S. K. Simić, , Encyclopedia of Mathematics and its Applications 66, Cambridge University Press, Cambridge, 1997
- [11] Secure and Non-Blind Watermarking Scheme for Color Images Based on DWT. M.G. Darwish, H.M. Onsi, M.Samy, G.S.El-Taweel. ICGST-GVIP Journal, Volume 5, Issue4, April 2005.
- [12] Dictionary of distances. Deza E. and Deza M. ELSEVIER, Chapter 21.
- [13] Visión por Computador. De la Cruz García J. M.; Pajares Martinsanz G. 2nd ed. Alfaomega (2002).
- [14] A Wreath Product Group Approach to Signal and Image Processing: Part I - Multiresolution Analysis. Foote, R. Mirchandani, G. Rockmore D. Healy D. Olson, T. IEEE transactions on signal processing, 2000, vol. 48, no1, pp. 102-132.
- [15] A Wreath Product Group Approach to Signal and Image Processing: Part II - Convolution, Correlation, and Applications. Foote, R. Mirchandani, G. Rockmore D. Healy D. Olson, T. IEEE transactions on signal processing, 2000, vol. 48, no1, pp. 102-132.
- [16] Signal compression based on models of human perception. N. Jayant, J. Johnston and R. Safranek, Proc. IEEE, vol. 81, no. 10, pp. 1385-1422, Oct. 1993.
- [17] Discrete Walsh-Hadamard Transform in Signal Processing. A.A.C.A.Jayathilake, A.A.I.Perera , M.A.P.Chamikara IJRIT Volume 1, Issue 1, January 2013, Pg. 80-89
- [18] Fast fractal image compression using the Hadamard transform. Jeng J.H.; Truong T.K.; Sheu J.R.; IEE Proceedings - Vision, Image and Signal Processing, Volume 147, Issue 6, December 2000.
- [19] Low Computational Complexity Variable Block Size (VBS) Partitioning for Motion Estimation using the Walsh Hadamard Transform (WHT). Chanyul Kim and Noel E.OConnor.Multimedia, 2009. ISM '09. 11th IEEE International Symposium on, 124 - 129, Dec. 2009.
- [20] The digital signal processing handbook. Madisetti V.; Williams D. CRC PRESS, IEEE PRESS, (1998).
- [21] Linear Ciphers and Random Sequence Generators with Multiple Clocks. Massey J.L. and Rueppel R.; Advances in Cryptology:Proc. Eurocrypt 84, Lecture Notes in Computer Science, vol. 209, Springer-Verlag, Berlin, 1984, pp. 74-87.
- [22] On the Optimality of SAFER+ Diffusion. Massey J.L.; Second Advanced Encryption Standard Candidate Conference (AES2), Rome, Italy, March 22-23, (1999).
- [23] SAFER K-64: A Byte- Oriented Block Ciphering Algorithm. Massey J. L.; Fast Software Encryption, Proceedings of the Cambridge Security Workshop, Cambridge, U.K, pp. 1- 17. (1998).
- [24] Image Compression: A Survey. Mehwish Rehman, Muhammad Sharif and Mudassar Raza. Research Journal of Applied Sciences, Engineering and Technology 7(4): 656-672, 2014.
- [25] One-Dimensional Hadamard Naturalness - Preserving Transform Reconstruction of Signals Degraded by Nonstationary Noise Processes. Adewole O. Osinubi and Robert A. King.IEEE TRANSACTIONS ON SIGNAL PROCESSING. VOL. 40, NO. 3. MARCH 1992.
- [26] Eigenspaces of Hamming graphs and unitary Cayley graphs. Torsten Sander. ARS MATHEMATICA CONTEMPORANEA 3 (2010) 13-19.
- [27] Performance Evaluation of VHDL Implementation of Proposed SAFER+ Security Algorithm and Pipelined AES Security Algorithm for Bluetooth Security Systems . Sharmila D.



Neelaveni R. ICGST-CNIR Journal, Volume 9, Issue 1, July 2009.

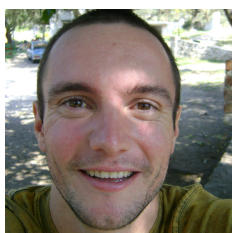
- [28] Compression Of Hyperspectral Image Using Discrete Wavelet Transform And Walsh Hadamard Transform. Sujithra D.S.; Manickam T.; Sudheer D.S.; International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 2, Issue 3, March 2013
- [29] Hadamard-Based Image Decomposition and Compression. Iren Valova and Yukio Kosugi. IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, VOL. 4, NO. 4, DECEMBER 2000.
- [30] Single-Scan 2D Hadamard NMR Spectroscopy. Assaf Tal, Boaz Shapira, and Lucio Frydman. Angew. Chem. Int. Ed. 2009, 48, pp 2732-2736
- [31] A naturalness-preserving transform for image coding and reconstruction. Yarlagadda R. and Hershey J. IEEE Trans. Acoust, Speech, Signal Processing, vol. ASSP-33, N^o 4. pp. 1005-1012, Aug. 1985.
- [32] Image Quality Assessment: From Error Visibility to Structural Similarity. Wang Z., Bovik A. C., Sheikh H. R. and Simoncelli E. P. IEEE Transactions on image Processing, VOL. 13, N^o 4, APRIL 2004.
- [33] <http://www.hlevkin.com/default.html#testimages>



Biographies



Fernando Levstein received his B.S. degree from the Universidad Nacional de Córdoba (UNC) in Argentina and a Ph.D. in Mathematics from the Massachusetts Institute of Technology. He is currently a Professor at the School of Mathematics, Astronomy and Physics (Fa.M.A.F) of the UNC. His research interests include Lie Theory, Harmonic Analysis and Algebraic Combinatorics.



Javier Lezama received a Ph. D. in Mathematics from the Universidad Nacional de Córdoba (UNC), Fa.M.A.F, in 2013. His research interests include algebraic combinatorics, mathematical digital image processing and image compression. He currently holds a fellowship of the National Council of Scientific and Technological Research (CONICET).



Ana Carolina Maldonado is a professor of the Department of Mathematics of the School of Exact, Physical and Natural Sciences (FCEPyN) of the UNC and is an assistant research of the National Council of Scientific and Technological Research (CONICET). She received a Ph. D. in Mathematics from the Universidad Nacional de Córdoba (UNC), FA.M.A.F, in 2003. Her research interest include algebraic combinatorics, mathematical morphology applied to digital image processing and image compression.



Daniel Penazzi is a professor of Fa.M.A.F, Universidad Nacional de Córdoba (UNC). He received a Ph D in Mathematics from the University of Minnesota in 1993. His research interest include algebraic combinatorics, cryptography, image compression, and topological dynamics.



Mauro Schilman received a B.S. degree in Computer Science from the Universidad Nacional de Córdoba (UNC), Fa.M.A.F, in 2015. His research interests include computer graphics, logic, bioinformatics and artificial intelligence.



6 Appendix: Examples of Compression

Here we give the results of our $d = 6$ algorithm with $c = 0.99$ using bzip2 as the lossless part. For fair comparison, we also include the results of using bzip2 alone (without our algorithm) in the first column. In the tables the numbers in the first row are the compressed ratios and the number below that are the sizes of the compressed image, in KB. The size of the original image is also displayed. The unit in the PSNR (Peak Signal to Noise Ratio) is dB. Remember that the higher is the PSNR, better is the quality of the reconstruction. (∞ is lossless). Usual values in image compression are between 30dB and 50dB.

Image	bzip2	d6bzip2
<i>Cameraman</i> (257KB)	5.5 : 1(46) PSNR : 31.74dB	18.3 : 1(14) PSNR : 31.74dB
<i>Finger</i> (257KB)	4.3 : 1(59) PSNR : 29.06dB	9.5 : 1(27) PSNR : 29.06dB
<i>Fingerprint</i> (257KB)	8.8 : 1(29) PSNR : 33.53dB	13.5 : 1(19) PSNR : 33.53dB
<i>Boats</i> (406KB)	1.5 : 1(254) PSNR : 34.33dB	9.0 : 1(45) PSNR : 34.33dB
<i>Zelda2</i> (406KB)	1.6 : 1(248) PSNR : 36.14dB	14.5 : 1(28) PSNR : 36.14dB
<i>Ai field2</i> (1000KB)	1.3 : 1(757) PSNR : 31.95dB	5.2 : 1(190) PSNR : 31.95dB
<i>AirplaneU2</i> (1000KB)	1.6 : 1(616) PSNR : 34.36dB	13.1 : 1(76) PSNR : 34.36dB
<i>Man</i> (1000KB)	1.4 : 1(706) PSNR : 33.24dB	6.8 : 1(147) PSNR : 33.24dB
<i>Pattern</i> (1000KB)	17.5 : 1(57) PSNR : 41.20dB	43.4 : 1(23) PSNR : 41.20dB
<i>Lena</i> (257KB)	1.5 : 1(171) PSNR : 34.53dB	8.2 : 1(31) PSNR : 34.53dB
<i>Baboon</i> (257KB)	1.2 : 1(213) PSNR : 30.74dB	3.3 : 1(77) PSNR : 30.74dB
<i>Aerial</i> (257KB)	1.3 : 1(187) PSNR : 30.06dB	4.3 : 1(59) PSNR : 30.06dB
<i>Ai field</i> (257KB)	1.3 : 1(185) PSNR : 30.34dB	4.0 : 1(63) PSNR : 30.34dB
<i>Bridge</i> (257KB)	1.8 : 1(139) PSNR : 31.06dB	3.6 : 1(70) PSNR : 31.06dB
<i>Clown</i> (257KB)	2.5 : 1(100) PSNR : 34.72dB	7.7 : 1(33) PSNR : 34.72dB
<i>Couple</i> (257KB)	1.4 : 1(173) PSNR : 32.91dB	6.2 : 1(41) PSNR : 32.91dB
<i>Crowd</i> (257KB)	1.6 : 1(152) PSNR : 33.52dB	5.5 : 1(46) PSNR : 33.52dB

<i>Dollar</i> (257KB)	1.2 : 1(203) PSNR : 30.41dB	3.2 : 1(78) PSNR : 30.41dB
<i>Girlface</i> (257KB)	1.8 : 1(139) PSNR : 35.74dB	9.8 : 1(26) PSNR : 35.74dB
<i>Houses</i> (257KB)	1.2 : 1(202) PSNR : 31.35dB	3.7 : 1(68) PSNR : 31.35dB
<i>Kiel</i> (257KB)	1 : 4 : 1(174) PSNR : 32.07dB	4 : 8 : 1(53) PSNR : 32.07dB
<i>Lighthouse</i> (257KB)	1.3 : 1(190) PSNR : 31.87dB	5.1 : 1(50) PSNR : 31.87dB
<i>Tank</i> (257KB)	1.7 : 1(145) PSNR : 32.18dB	9.1 : 1(28) PSNR : 32.18dB
<i>Tank2</i> (257KB)	1.5 : 1(171) PSNR : 31.22dB	5.9 : 1(43) PSNR : 31.22dB
<i>Truck</i> (257KB)	1.7 : 1(147) PSNR : 32.18dB	8.2 : 1(31) PSNR : 32.18dB
<i>Trucks</i> (257KB)	1.5 : 1(171) PSNR : 30.90dB	5.2 : 1(49) PSNR : 30.90dB
<i>Zelda</i> (257KB)	2.6 : 1(96) PSNR : 35.69dB	11.6 : 1(22) PSNR : 35.69dB



